

# Technical Report # KU-EC-10-2

## The Distribution of the Domination Number of a Family of Random Interval Catch Digraphs

Elvan Ceyhan\*

March 30, 2010

### Abstract

We study a new kind of proximity graphs called proportional-edge proximity catch digraphs (PCDs) in a randomized setting. PCDs are a special kind of random catch digraphs that have been developed recently and have applications in statistical pattern classification and spatial point pattern analysis. PCDs are also a special type of intersection digraphs; and for one-dimensional data, the proportional-edge PCD family is also a family of random interval catch digraphs. We present the exact (and asymptotic) distribution of the domination number of this PCD family for uniform (and non-uniform) data in one dimension. We also provide several extensions of this random catch digraph by relaxing the expansion and centrality parameters, thereby determine the parameters for which the asymptotic distribution is non-degenerate. We observe sudden jumps (from degeneracy to non-degeneracy or from a non-degenerate distribution to another) in the asymptotic distribution of the domination number at certain parameter values.

*Keywords:* asymptotic distribution; class cover catch digraph; degenerate distribution; exact distribution; intersection digraph; proximity catch digraph; proximity map; random graph; uniform distribution

*AMS 2000 Subject Classification:* 05C80; 05C20; 60D05; 60C05; 62E20

---

\*Address: Department of Mathematics, Koç University, 34450 Sarıyer, Istanbul, Turkey. e-mail: elceyhan@ku.edu.tr, tel: +90 (212) 338-1845, fax: +90 (212) 338-1559.

# 1 Introduction

The proximity catch digraphs (PCDs) were motivated by their applications in pattern classification and spatial pattern analysis, hence they have become focus of considerable attention recently. The PCDs are a special type of proximity graphs which were introduced by Toussaint (1980). A *digraph* is a directed graph with vertex set  $V$  and arcs (directed edges) each of which is from one vertex to another based on a binary relation. Then the pair  $(p, q) \in V \times V$  is an ordered pair which stands for an *arc* from vertex  $p$  to vertex  $q$  in  $V$ . For example, *nearest neighbor (di)graph* which is defined by placing an arc between each vertex and its nearest neighbor is a proximity digraph where vertices represent points in some metric space (Paterson and Yao (1992)). PCDs are *data-random digraphs* in which each vertex corresponds to a data point and arcs are defined in terms of some bivariate relation on the data.

The PCDs are closely related to the class cover problem of Cannon and Cowen (2000). Let  $(\Omega, \mathcal{M})$  be a measurable space and  $\mathcal{X}_n = \{X_1, X_2, \dots, X_n\}$  and  $\mathcal{Y}_m = \{Y_1, Y_2, \dots, Y_m\}$  be two sets of  $\Omega$ -valued random variables from classes  $\mathcal{X}$  and  $\mathcal{Y}$ , respectively, with joint probability distribution  $F_{X,Y}$ . Let  $d(\cdot, \cdot) : \Omega \times \Omega \rightarrow [0, \infty)$  be any distance function. The *class cover problem* for a target class, say  $\mathcal{X}$ , refers to finding a collection of neighborhoods,  $N_i$  around  $X_i$  such that (i)  $\mathcal{X}_n \subseteq (\cup_i N_i)$  and (ii)  $\mathcal{Y}_m \cap (\cup_i N_i) = \emptyset$ . A collection of neighborhoods satisfying both conditions is called a *class cover*. A cover satisfying (i) is a *proper cover* of  $\mathcal{X}_n$  while a cover satisfying (ii) is a *pure cover* relative to  $\mathcal{Y}_m$ . This article is on the *cardinality of smallest class covers*; that is, class covers satisfying both (i) and (ii) with the smallest number of neighborhoods. See Cannon and Cowen (2000) and Priebe et al. (2001) for more on the class cover problem.

The first type of PCD was class cover catch digraph (CCCD) introduced by Priebe et al. (2001) who gave the exact distribution of its domination number for uniform data from two classes in  $\mathbb{R}$ . DeVinney et al. (2002), Marchette and Priebe (2003), Priebe et al. (2003a), Priebe et al. (2003b), DeVinney and Priebe (2006) extended the CCCDs to higher dimensions and demonstrated that CCCDs are a competitive alternative to the existing methods in classification. Furthermore, DeVinney and Wierman (2003) proved a SLLN result for the one-dimensional class cover problem; Wierman and Xiang (2008) provided a generalized SLLN result and Xiang and Wierman (2009) provided a CLT result for CCCD based on one-dimensional data. However, CCCDs have some disadvantages in higher dimensions; namely, finding the minimum dominating set for CCCDs is an NP-hard problem in general, although a simple linear time algorithm is available for one dimensional data (Priebe et al. (2001)); and the exact and the asymptotic distributions of the domination number of the CCCDs are not analytically tractable in multiple dimensions. Ceyhan and Priebe (2003, 2005) introduced the central similarity proximity maps and proportional-edge proximity maps for data in  $\mathbb{R}^d$  with  $d > 1$  and the associated random PCDs with the purpose of avoiding the above mentioned problems. The asymptotic distribution of the domination number of the proportional-edge PCD is calculated for data in  $\mathbb{R}^2$  and then the domination number is used as a statistic for testing bivariate spatial patterns (Ceyhan and Priebe (2005), Ceyhan (2010)). The relative density of these two PCD families is also calculated and used for the same purpose (Ceyhan et al. (2006) and Ceyhan et al. (2007)). Moreover, the distribution of the domination number of CCCDs is derived for non-uniform data (Ceyhan (2008)).

In this article, we provide the exact (and asymptotic) distribution of the domination number of proportional-edge PCDs for uniform (and non-uniform) one-dimensional data. First, some special cases and bounds for the domination number of proportional-edge PCDs is presented, then the domination number is investigated for uniform data in one interval (in  $\mathbb{R}$ ) and the analysis is generalized to uniform data in multiple intervals and to non-uniform data in one and multiple intervals. These results can be seen as generalizations of the results of Ceyhan (2008). Some trivial proofs are omitted, shorter proofs are given in the main body of the article; while longer proofs are deferred to the Appendix.

We define the proportional-edge PCDs and their domination number in Section 2, provide the exact and asymptotic distributions of the domination number of proportional-edge PCDs for uniform data in one interval in Section 3, discuss the distribution of the domination number for data from a general distribution in Section 4. We extend these results to multiple intervals in Section 5, and provide discussion and conclusions in Section 6. For convenience in notation and presentation, we resort to non-standard extended (perhaps abused) forms of

Bernoulli and Binomial distributions, denoted  $\text{BER}(p)$  and  $\text{BIN}(n, p)$ , respectively, where  $p$  is the probability of success and  $n$  is the number of trials. Throughout the article, we take  $p \in [0, 1]$  (unlike  $p \in (0, 1)$ ) and if  $X \sim \text{BER}(p)$ , then  $P(X = 1) = p$  and  $P(X = 0) = 1 - p$ . If  $Y \sim \text{BIN}(n, p)$ , then  $\frac{P(Y=k)}{p^k(1-p)^{n-k}} = \binom{n}{k}$  for  $p \in (0, 1)$  and  $k \in \{0, 1, 2, \dots, n\}$  and  $P(Y = n) = 1$  for  $p = 1$  and  $P(Y = 0) = 1$  for  $p = 0$ .

## 2 Proportional-Edge Proximity Catch Digraphs

Consider the map  $N : \Omega \rightarrow \wp(\Omega)$  where  $\wp(\Omega)$  represents the power set of  $\Omega$ . Then given  $\mathcal{Y}_m \subseteq \Omega$ , the *proximity map*  $N(\cdot) : \Omega \rightarrow \wp(\Omega)$  associates with each point  $x \in \Omega$  a *proximity region*  $N(x) \subseteq \Omega$ . For  $B \subseteq \Omega$ , the  $\Gamma_1$ -region is the image of the map  $\Gamma_1(\cdot, N) : \wp(\Omega) \rightarrow \wp(\Omega)$  that associates the region  $\Gamma_1(B, N) := \{z \in \Omega : B \subseteq N(z)\}$  with the set  $B$ . For a point  $x \in \Omega$ , we denote  $\Gamma_1(\{x\}, N)$  as  $\Gamma_1(x, N)$ . Notice that while the proximity region is defined for one point, a  $\Gamma_1$ -region is defined for a set of points. The *data-random PCD* has the vertex set  $\mathcal{V} = \mathcal{X}_n$  and arc set  $\mathcal{A}$  defined by  $(X_i, X_j) \in \mathcal{A}$  iff  $X_j \in N(X_i)$ .

Let  $\Omega = \mathbb{R}$  and  $Y_{(i)}$  be the  $i^{\text{th}}$  order statistic (i.e.,  $i^{\text{th}}$  smallest value) of  $\mathcal{Y}_m$  for  $i = 1, 2, \dots, m$  with the additional notation for  $i \in \{0, m+1\}$  as

$$-\infty =: Y_{(0)} < Y_{(1)} < \dots < Y_{(m)} < Y_{(m+1)} := \infty.$$

Then  $Y_{(i)}$  partition  $\mathbb{R}$  into  $(m+1)$  intervals which is called the *intervalization* of  $\mathbb{R}$  by  $\mathcal{Y}_m$ . Let also that  $\mathcal{I}_i := (Y_{(i)}, Y_{(i+1)})$  for  $i \in \{0, 1, 2, \dots, m\}$  and  $M_{c,i} := Y_{(i)} + c(Y_{(i+1)} - Y_{(i)})$  (i.e.,  $M_{c,i} \in \mathcal{I}_i$  such that  $c \times 100\%$  of length of  $\mathcal{I}_i$  is to the left of  $M_{c,i}$ ). We define the proportional-edge proximity region with the expansion parameter  $r \geq 1$  and centrality parameter  $c \in [0, 1]$  for two one-dimensional data sets,  $\mathcal{X}_n$  and  $\mathcal{Y}_m$ , from classes  $\mathcal{X}$  and  $\mathcal{Y}$ , respectively, as follows. For  $x \in \mathcal{I}_i$  with  $i \in \{1, 2, \dots, m-1\}$

$$N(x, r, c) = \begin{cases} (Y_{(i)}, Y_{(i)} + r(x - Y_{(i)})) \cap \mathcal{I}_i & \text{if } x \in (Y_{(i)}, M_{c,i}), \\ (Y_{(i+1)} - r(Y_{(i+1)} - x), Y_{(i+1)}) \cap \mathcal{I}_i & \text{if } x \in (M_{c,i}, Y_{(i+1)}), \end{cases} \quad (1)$$

Additionally, for  $x \in \mathcal{I}_i$  with  $i \in \{0, m\}$

$$N(x, r, c) = \begin{cases} (Y_{(1)} - r(Y_{(1)} - x), Y_{(1)}) & \text{if } x < Y_{(1)}, \\ (Y_{(m)}, Y_{(m)} + r(x - Y_{(m)})) & \text{if } x > Y_{(m)}. \end{cases} \quad (2)$$

Notice that for  $i \in \{0, m\}$ , the proportional-edge proximity region does not depend on the centrality parameter  $c$ . For  $x \in \mathcal{Y}_m$ , we define  $N(x, r, c) = \{x\}$  for all  $r \geq 1$  and if  $x = M_{c,i}$ , then in Equation (1), we arbitrarily assign  $N(x, r, c)$  to be one of  $(Y_{(i)}, Y_{(i)} + r(x - Y_{(i)})) \cap \mathcal{I}_i$  or  $(Y_{(i+1)} - r(Y_{(i+1)} - x), Y_{(i+1)}) \cap \mathcal{I}_i$ . For  $c = 0$ , we have  $(M_{c,i}, Y_{(i+1)}) = \mathcal{I}_i$  and for  $c = 1$ , we have  $(Y_{(i)}, M_{c,i}) = \mathcal{I}_i$ . So, we set  $N(x, r, 0) := (Y_{(i+1)} - r(Y_{(i+1)} - x), Y_{(i+1)}) \cap \mathcal{I}_i$  and  $N(x, r, 1) := (Y_{(i)}, Y_{(i)} + r(x - Y_{(i)})) \cap \mathcal{I}_i$ . For  $r > 1$ , we have  $x \in N(x, r, c)$  for all  $x \in \mathcal{I}_i$ . Furthermore,  $\lim_{r \rightarrow \infty} N(x, r, c) = \mathcal{I}_i$  for all  $x \in \mathcal{I}_i$ , so we define  $N(x, \infty, c) = \mathcal{I}_i$  for all such  $x$ .

For  $X_i \stackrel{iid}{\sim} F$ , with the additional assumption that the non-degenerate one-dimensional probability density function (pdf)  $f$  exists with support  $\mathcal{S}(F) \subseteq \mathcal{I}_i$  and  $f$  is continuous around  $M_{c,i}$  and around the end points of  $\mathcal{I}_i$ , implies that the special cases in the construction of  $N(\cdot, r, c) - X$  falls at  $M_{c,i}$  or the end points of  $\mathcal{I}_i$  — occurs with probability zero. For such an  $F$ , the region  $N(X_i, r, c)$  is an interval a.s.

The data-random proportional-edge PCD has the vertex set  $\mathcal{X}_n$  and arc set  $\mathcal{A}$  defined by  $(X_i, X_j) \in \mathcal{A}$  iff  $X_j \in N(X_i, r, c)$ . We call such digraphs  $\mathcal{D}_{n,m}(r, c)$ -digraphs. A  $\mathcal{D}_{n,m}(r, c)$ -digraph is a *pseudo digraph* according some authors, if loops are allowed (see, e.g., Chartrand and Lesniak (1996)). The  $\mathcal{D}_{n,m}(r, c)$ -digraphs are closely related to the *proximity graphs* of Jaromczyk and Toussaint (1992) and might be considered as a special case of *covering sets* of Tuza (1994) and *intersection digraphs* of Sen et al. (1989). Our data-random proximity digraph is a *vertex-random digraph* and is not a standard random graph (see, e.g., Janson et al.

(2000)). The randomness of a  $\mathcal{D}_{n,m}(r, c)$ -digraph lies in the fact that the vertices are random with the joint distribution  $F_{X,Y}$ , but arcs  $(X_i, X_j)$  are deterministic functions of the random variable  $X_j$  and the random set  $N(X_i, r, c)$ . In  $\mathbb{R}$ , the data-random PCD is a special case of *interval catch digraphs* (see, e.g., Sen et al. (1989) and Prisner (1994)). Furthermore, when  $r = 2$  and  $c = 1/2$  (i.e.,  $M_{c,i} = (Y_{(i)} + Y_{(i+1)})/2$ ) we have  $N(x, r, c) = B(x, r(x))$  where  $B(x, r(x))$  is the ball centered at  $x$  with radius  $r(x) = d(x, \mathcal{Y}_m) = \min_{y \in \mathcal{Y}_m} d(x, y)$ . The region  $N(x, 2, 1/2)$  corresponds to the proximity region which gives rise to the CCCD of Priebe et al. (2001).

## 2.1 Domination Number of Random $\mathcal{D}_{n,m}(r, c)$ -digraphs

In a digraph  $D = (\mathcal{V}, \mathcal{A})$  of order  $|\mathcal{V}| = n$ , a vertex  $v$  *dominates* itself and all vertices of the form  $\{u : (v, u) \in \mathcal{A}\}$ . A *dominating set*,  $S_D$ , for the digraph  $D$  is a subset of  $\mathcal{V}$  such that each vertex  $v \in \mathcal{V}$  is dominated by a vertex in  $S_D$ . A *minimum dominating set*,  $S_D^*$ , is a dominating set of minimum cardinality; and the *domination number*, denoted  $\gamma(D)$ , is defined as  $\gamma(D) := |S_D^*|$ , where  $|\cdot|$  is the set cardinality functional (West (2001)). If a minimum dominating set consists of only one vertex, we call that vertex a *dominating vertex*. The vertex set  $\mathcal{V}$  itself is always a dominating set, so  $\gamma(D) \leq n$ .

Let  $\mathcal{F}(\mathbb{R}^d) := \{F_{X,Y} \text{ on } \mathbb{R}^d \text{ with } P(X = Y) = 0\}$ . As in Priebe et al. (2001) and Ceyhan (2008), we consider  $\mathcal{D}_{n,m}(r, c)$ -digraphs for which  $\mathcal{X}_n$  and  $\mathcal{Y}_m$  are random samples from  $F_X$  and  $F_Y$ , respectively, and the joint distribution of  $X, Y$  is  $F_{X,Y} \in \mathcal{F}(\mathbb{R}^d)$ . We call such digraphs  $\mathcal{F}(\mathbb{R}^d)$ -random  $\mathcal{D}_{n,m}(r, c)$ -digraphs and focus on the random variable  $\gamma(D)$ . To make the dependence on sample sizes  $n$  and  $m$ , the distribution  $F$ , and the parameters  $r$  and  $c$  explicit, we use  $\gamma_{n,m}(F, r, c)$  instead of  $\gamma(D)$ . For  $n \geq 1$  and  $m \geq 1$ , it is trivial to see that  $1 \leq \gamma_{n,m}(F, r, c) \leq n$ , and  $1 \leq \gamma_{n,m}(F, r, c) < n$  for nontrivial digraphs.

## 2.2 Special Cases for the Distribution of the Domination Number of $\mathcal{F}(\mathbb{R})$ -random $\mathcal{D}_{n,m}(r, c)$ -digraphs

Let  $\mathcal{X}_n$  and  $\mathcal{Y}_m$  be two samples from  $\mathcal{F}(\mathbb{R})$ ,  $\mathcal{X}_{[i]} := \mathcal{X}_n \cap \mathcal{I}_i$ , and  $\mathcal{Y}_{[i]} := \{Y_{(i)}, Y_{(i+1)}\}$  for  $i = 0, 1, 2, \dots, m$ . This yields a disconnected digraph with subdigraphs each of which might be null or itself disconnected. Let  $D_{[i]}$  be the component of the random  $\mathcal{D}_{n,m}(r, c)$ -digraph induced by the pair  $\mathcal{X}_{[i]}$  and  $\mathcal{Y}_{[i]}$  for  $i = 0, 1, 2, \dots, m$ ,  $n_i := |\mathcal{X}_{[i]}|$ , and  $F_i$  be the density  $F_X$  restricted to  $\mathcal{I}_i$ , and  $\gamma_{n_i,2}(F_i, r, c)$  be the domination number of  $D_{[i]}$ . Let also that  $M_{c,i} \in \mathcal{I}_i$  be the point that divides the interval  $\mathcal{I}_i$  in ratios  $c$  and  $1 - c$  (i.e., length of the subinterval to the left of  $M_{c,i}$  is  $c \times 100\%$  of the length of  $\mathcal{I}_i$ ). Then  $\gamma_{n,m}(F, r, c) = \sum_{i=0}^m (\gamma_{n_i,2}(F_i, r, c) \mathbf{I}(n_i > 0))$  where  $\mathbf{I}(\cdot)$  is the indicator function. We study the simpler random variable  $\gamma_{n_i,2}(F_i, r, c)$  first. The following lemma follows trivially.

**Lemma 2.1.** *For  $i \in \{0, m\}$ , we have  $\gamma_{n_i,2}(F_i, r, c) = \mathbf{I}(n_i > 0)$  for all  $r \geq 1$ .*

Let  $\Gamma_1(B, r, c)$  be the  $\Gamma_1$ -region for set  $B$  associated with the proximity map on  $N(\cdot, r, c)$ .

**Lemma 2.2.** *The  $\Gamma_1$ -region for  $\mathcal{X}_{[i]}$  in  $\mathcal{I}_i$  with  $r \geq 1$  and  $c \in [0, 1]$  is*

$$\Gamma_1(\mathcal{X}_{[i]}, r, c) = \left( Y_{(i)} + \frac{\max(\mathcal{X}_{[i]})}{r}, M_{c,i} \right] \cup \left[ M_{c,i}, Y_{(i+1)} - \frac{Y_{(i+1)} - \min(\mathcal{X}_{[i]})}{r} \right)$$

*with the understanding that the intervals  $(a, b)$ ,  $(a, b]$ , and  $[a, b)$  are empty if  $a \geq b$ .*

**Proof:** By definition,  $\Gamma_1(\mathcal{X}_{[i]}, r, c) = \{x \in \mathcal{I}_i : \mathcal{X}_{[i]} \subset N(x, r, c)\}$ . Suppose  $r \geq 1$  and  $c \in [0, 1]$ . Then for  $x \in (Y_{(i)}, M_{c,i}]$ , we have  $\mathcal{X}_{[i]} \subset N(x, r, c)$  iff  $Y_{(i)} + r(x - Y_{(i)}) > \max(\mathcal{X}_{[i]})$  iff  $x > Y_{(i)} + \frac{\max(\mathcal{X}_{[i]})}{r}$ . Likewise for  $x \in [M_{c,i}, Y_{(i+1)})$ , we have  $\mathcal{X}_{[i]} \subset N(x, r, c)$  iff  $Y_{(i+1)} - r(Y_{(i+1)} - x) < \min(\mathcal{X}_{[i]})$  iff  $x < Y_{(i+1)} - \frac{Y_{(i+1)} - \min(\mathcal{X}_{[i]})}{r}$ . Therefore  $\Gamma_1(\mathcal{X}_{[i]}, r, c) = \left( Y_{(i)} + \frac{\max(\mathcal{X}_{[i]})}{r}, M_{c,i} \right] \cup \left[ M_{c,i}, Y_{(i+1)} - \frac{Y_{(i+1)} - \min(\mathcal{X}_{[i]})}{r} \right)$ . ■

Notice that if  $\mathcal{X}_{[i]} \cap \Gamma_1(\mathcal{X}_{[i]}, r, c) \neq \emptyset$ , we have  $\gamma_{n_i,2}(F_i, r, c) = 1$ , hence the name  $\Gamma_1$ -region and the notation  $\Gamma_1(\cdot)$ . For  $i = 1, 2, 3, \dots, (m-1)$  and  $n_i > 0$ , we prove that  $\gamma_{n_i,2}(F_i, r, c) = 1$  or 2 with distribution dependent probabilities. Hence, to find the distribution of  $\gamma_{n_i,2}(F_i, r, c)$ , it suffices to find  $P(\gamma_{n_i,2}(F_i, r, c) = 1)$  or  $p_{n_i}(F_i, r, c) := P(\gamma_{n_i,2}(F_i, r, c) = 2)$ . For computational convenience, we employ the latter in our calculations, henceforth.

**Theorem 2.3.** *For  $i = 1, 2, 3, \dots, (m-1)$ , let the support of  $F_i$  have a positive Lebesgue measure. Then for  $n_i > 1$ ,  $r \in (1, \infty)$ , and  $c \in (0, 1)$ , we have  $\gamma_{n_i,2}(F_i, r, c) \sim 1 + \text{BER}(p_{n_i}(F_i, r, c))$ . Furthermore,  $\gamma_{1,2}(F_i, r, c) = 1$  for all  $r \geq 1$  and  $c \in [0, 1]$ ;  $\gamma_{n_i,2}(F_i, r, 0) = \gamma_{n_i,2}(F_i, r, 1) = 1$  for all  $n_i \geq 1$  and  $r \geq 1$ ; and  $\gamma_{n_i,2}(F_i, \infty, c) = 1$  for all  $n_i \geq 1$  and  $c \in [0, 1]$ .*

**Proof:** Let  $X_i^- := \operatorname{argmin}_{x \in \mathcal{X}_{[i]} \cap (Y_{(i)}, M_{c,i})} d(x, M_{c,i})$  provided that  $\mathcal{X}_{[i]} \cap (Y_{(i)}, M_{c,i}) \neq \emptyset$ , and  $X_i^+ := \operatorname{argmin}_{x \in \mathcal{X}_{[i]} \cap (M_{c,i}, Y_{(i+1)})} d(x, M_{c,i})$  provided that  $\mathcal{X}_{[i]} \cap (M_{c,i}, Y_{(i+1)}) \neq \emptyset$ . That is,  $X_i^-$  and  $X_i^+$  are closest class  $\mathcal{X}$  points (if they exist) to  $M_{c,i}$  from left and right, respectively. Notice that since  $n_i > 0$ , at least one of  $X_i^-$  and  $X_i^+$  exists a.s. If  $\mathcal{X}_{[i]} \cap (Y_{(i)}, M_{c,i}) = \emptyset$ , then  $\mathcal{X}_{[i]} \subset N(X_i^+, r, c)$ ; so  $\gamma_{n_i,2}(F_i, r, c) = 1$ . Similarly, if  $\mathcal{X}_{[i]} \cap (M_{c,i}, Y_{(i)}) = \emptyset$ , then  $\mathcal{X}_{[i]} \subset N(X_i^-, r, c)$ ; so  $\gamma_{n_i,2}(F_i, r, c) = 1$ . If both of  $\mathcal{X}_{[i]} \cap (Y_{(i)}, M_{c,i})$  and  $\mathcal{X}_{[i]} \cap (M_{c,i}, Y_{(i)})$  are nonempty, then  $\mathcal{X}_{[i]} \subset N(X_i^-, r, c) \cup N(X_i^+, r, c)$ , so  $\gamma_{n_i,2}(F_i, r, c) \leq 2$ . Since  $n_i > 0$ , we have  $1 \leq \gamma_{n_i,2}(F_i, r, c) \leq 2$ . The desired result follows, since the probabilities  $1 - p_{n_i}(F, r, c) = P(\gamma_{n_i,2}(F_i, r, c) = 1)$  and  $p_{n_i}(F, r, c) = P(\gamma_{n_i,2}(F_i, r, c) = 2)$  are both positive. The special cases in the theorem follow by construction. ■

The probability  $p_{n_i}(F, r, c) = P(\mathcal{X}_{[i]} \cap \Gamma_1(\mathcal{X}_{[i]}, r, c) = \emptyset)$  depends on the conditional distribution  $F_{X|Y}$  and the interval  $\Gamma_1(\mathcal{X}_{[i]}, r, c)$ , which, if known, will make possible the calculation of  $p_{n_i}(F_i, r, c)$ . As an immediate result of Lemma 2.1 and Theorem 2.3, we have the following upper bound for  $\gamma_{n,m}(F, r, c)$ .

**Theorem 2.4.** *Let  $D_{n,m}(r, c)$  be an  $\mathcal{F}(\mathbb{R})$ -random  $\mathcal{D}_{n,m}(r, c)$ -digraph and  $k_1, k_2$ , and  $k_3$  be three natural numbers defined as  $k_1 := \sum_{i=1}^{m-1} \mathbf{I}(n_i > 1)$ ,  $k_2 := \sum_{i=1}^{m-1} \mathbf{I}(n_i = 1)$ , and  $k_3 := \sum_{i \in \{0,m\}} \mathbf{I}(n_i > 0)$ . Then for  $n \geq 1$ ,  $m \geq 1$ ,  $r \geq 1$ , and  $c \in [0, 1]$ , we have  $1 \leq \gamma_{n,m}(F, r, c) \leq 2k_1 + k_2 + k_3 \leq \min(n, 2m)$ . Furthermore,  $\gamma_{1,m}(F, r, c) = 1$  for all  $m \geq 1$ ,  $r \geq 1$ , and  $c \in [0, 1]$ ;  $\gamma_{n,1}(F, r, c) = \sum_{i \in \{0,1\}} \mathbf{I}(n_i > 0)$  for all  $n \geq 1$  and  $r \geq 1$ ;  $\gamma_{1,1}(F, r, c) = 1$  for all  $r \geq 1$ ;  $\gamma_{n,m}(F, r, 0) = \gamma_{n,m}(F, r, 1) = k_1 + k_2 + k_3$  for all  $m > 1$ ,  $n \geq 1$ , and  $r \geq 1$ ; and  $\gamma_{n,m}(F, \infty, c) = k_1 + k_2 + k_3$  for all  $m > 1$ ,  $n \geq 1$ , and  $c \in [0, 1]$ .*

**Proof:** Suppose  $n \geq 1$ ,  $m \geq 1$ ,  $r \geq 1$ , and  $c \in [0, 1]$ . Then for  $i = 1, 2, \dots, (m-1)$ , by Theorem 2.3, we have  $\gamma_{n_i,2}(F_i, r, c) \in \{1, 2\}$  provided that  $n_i > 1$ , and  $\gamma_{1,2}(F_i, r, c) = 1$ . For  $i \in \{0, m\}$ , by Lemma 2.1, we have  $\gamma_{n_i,2}(F_i, r, c) = \mathbf{I}(n_i > 0)$ . Since  $\gamma_{n,m}(F, r, c) = \sum_{i=0}^m (\gamma_{n_i,2}(F_i, r, c) \mathbf{I}(n_i > 0))$ , the desired result follows. The special cases in the theorem follow by construction. ■

For  $r = 1$ , the distribution of  $\gamma_{n_i,2}(F_i, r, c)$  is simpler and the distribution of  $\gamma_{n,m}(F_i, r, c)$  has simpler upper bounds.

**Theorem 2.5.** *Let  $D_{n,m}(1, c)$  be an  $\mathcal{F}(\mathbb{R})$ -random  $\mathcal{D}_{n,m}(1, c)$ -digraph,  $k_3$  be defined as in Theorem 2.4, and  $k_4$  be a natural number defined as  $k_4 := \sum_{i=1}^{m-1} [\mathbf{I}(|\mathcal{X}_{[i]} \cap (Y_{(i)}, M_{c,i})| > 0) + \mathbf{I}(|\mathcal{X}_{[i]} \cap (M_{c,i}, Y_{(i+1)})| > 0)]$ . Then for  $n \geq 1$ ,  $m > 1$ , and  $c \in [0, 1]$ , we have  $1 \leq \gamma_{n,m}(F, 1, c) = k_3 + k_4 \leq \min(n, 2m)$ .*

**Proof:** Suppose  $n \geq 1$ ,  $m > 1$ , and  $c \in [0, 1]$  and let  $X_i^-$  and  $X_i^+$  be defined as in the proof of Theorem 2.3. Then by construction,  $\mathcal{X}_{[i]} \cap (Y_{(i)}, M_{c,i}) \subset N(X_i^-, 1, c)$ , but  $N(X_i^-, 1, c) \subseteq (Y_{(i)}, M_{c,i})$ . So  $[\mathcal{X}_{[i]} \cap (M_{c,i}, Y_{(i+1)})] \cap N(X_i^-, 1, c) = \emptyset$ . Similarly  $\mathcal{X}_{[i]} \cap (M_{c,i}, Y_{(i+1)}) \subset N(X_i^+, 1, c)$  and  $[\mathcal{X}_{[i]} \cap (Y_{(i)}, M_{c,i})] \cap N(X_i^+, 1, c) = \emptyset$ . Then  $\gamma_{n_i,2}(F_i, 1, c) = 1$ , if  $\mathcal{X}_{[i]} \subset (Y_{(i)}, M_{c,i})$  or  $\mathcal{X}_{[i]} \subset (M_{c,i}, Y_{(i+1)})$ , and  $\gamma_{n,m}(F, 1, c) = 2$ , if  $\mathcal{X}_{[i]} \cap (Y_{(i)}, M_{c,i}) \neq \emptyset$  and  $\mathcal{X}_{[i]} \cap (M_{c,i}, Y_{(i+1)}) \neq \emptyset$ . Hence for  $i = 1, 2, 3, \dots, (m-1)$ , we have  $\gamma_{n,m}(F, 1, c) = \mathbf{I}(|\mathcal{X}_{[i]} \cap (Y_{(i)}, M_{c,i})| > 0) + \mathbf{I}(|\mathcal{X}_{[i]} \cap (M_{c,i}, Y_{(i+1)})| > 0)$ , and for  $i \in \{0, m\}$ , we have  $\gamma_{n_i,2}(F_i, 1, c) = \mathbf{I}(n_i > 0)$ . Since  $\gamma_{n,m}(F, 1, c) = \sum_{i=0}^m (\gamma_{n_i,2}(F_i, 1, c) \mathbf{I}(n_i > 0))$ , the desired result follows. ■

Based on Theorem 2.5, we have  $P(\gamma_{n,i,2}(F, 1, c) = 1) = P(\mathcal{X}_{[i]} \subset (Y_{(i)}, M_{c,i})) + P(\mathcal{X}_{[i]} \subset (M_{c,i}, Y_{(i+1)}))$  and  $P(\gamma_{n,i,2}(F, 1, c) = 2) = P(\mathcal{X}_{[i]} \cap (Y_{(i)}, M_{c,i}) \neq \emptyset, \mathcal{X}_{[i]} \cap (M_{c,i}, Y_{(i+1)}) \neq \emptyset)$ .

### 3 The Distribution of the Domination Number of Proportional-Edge PCDs for Uniform Data in One Interval

In the special case of fixed  $\mathcal{Y}_2 = \{y_1, y_2\}$  with  $-\infty < y_1 < y_2 < \infty$  and  $\mathcal{X}_n = \{X_1, X_2, \dots, X_n\}$  a random sample from  $\mathcal{U}(y_1, y_2)$ , the uniform distribution on  $(y_1, y_2)$ , we have a  $\mathcal{D}_{n,2}(r, c)$ -digraph for which  $F_X = \mathcal{U}(y_1, y_2)$ . We call such digraphs as  $\mathcal{U}(y_1, y_2)$ -random  $\mathcal{D}_{n,2}(r, c)$ -digraphs and provide the exact distributions of their domination number for the whole range of  $r$  and  $c$ . Let  $\gamma_{n,2}(\mathcal{U}, r, c)$  be the domination number of the PCD based on  $N(\cdot, r, c)$  and  $\mathcal{X}_n$  and  $p_n(\mathcal{U}, r, c) := P(\gamma_{n,2}(\mathcal{U}, r, c) = 2)$ , and  $p(\mathcal{U}, r, c) := \lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, c)$ . We present a “scale invariance” result for  $N(\cdot, r, c)$ . This invariance property will simplify the notation and calculations in our subsequent analysis by allowing us to consider the special case of the unit interval,  $(0, 1)$ .

**Proposition 3.1.** (*Scale Invariance Property*) Suppose  $\mathcal{X}_n$  is a random sample (i.e., a set of iid random variables) from  $\mathcal{U}(y_1, y_2)$ . Then for any  $r \in [1, \infty]$  the distribution of  $\gamma_{n,2}(\mathcal{U}, r, c)$  is independent of  $\mathcal{Y}_2$  and hence the support interval  $(y_1, y_2)$ .

**Proof:** Let  $\mathcal{X}_n$  be a random sample from  $\mathcal{U}(y_1, y_2)$  distribution. Any  $\mathcal{U}(y_1, y_2)$  random variable can be transformed into a  $\mathcal{U}(0, 1)$  random variable by the transformation  $\phi(x) = (x - y_1)/(y_2 - y_1)$ , which maps intervals  $(t_1, t_2) \subseteq (y_1, y_2)$  to intervals  $(\phi(t_1), \phi(t_2)) \subseteq (0, 1)$ . That is, if  $X \sim \mathcal{U}(y_1, y_2)$ , then we have  $\phi(X) \sim \mathcal{U}(0, 1)$  and  $P(X \in (t_1, t_2)) = P(\phi(X) \in (\phi(t_1), \phi(t_2)))$  for all  $(t_1, t_2) \subseteq (y_1, y_2)$ . So, without loss of generality, we can assume  $\mathcal{X}_n$  is a random sample from the  $\mathcal{U}(0, 1)$  distribution. Therefore, the distribution of  $\gamma_{n,2}(\mathcal{U}, r, c)$  does not depend on the support interval  $(y_1, y_2)$ . ■

Note that scale invariance of  $\gamma_{n,2}(F, \infty, c)$  follows trivially for all  $\mathcal{X}_n$  from any  $F$  with support in  $(y_1, y_2)$ , since for  $r = \infty$ , we have  $\gamma_{n,2}(F, \infty, c) = 1$  a.s. for all  $n > 1$  and  $c \in (0, 1)$ . The scale invariance of  $\gamma_{1,2}(F, r, c)$  holds for  $n = 1$  for all  $r \geq 1$  and  $c \in [0, 1]$ , and scale invariance of  $\gamma_{n,2}(F, r, c)$  with  $c \in \{0, 1\}$  holds for  $n \geq 1$  and  $r \geq 1$ , as well. Based on Proposition 3.1, for uniform data, we may assume that  $(y_1, y_2)$  is the unit interval  $(0, 1)$  for  $N(\cdot, r, c)$  with general  $c$ . Then the proportional-edge proximity region for  $x \in (0, 1)$  with parameters  $r \geq 1$  and  $c \in [0, 1]$  becomes

$$N(x, r, c) = \begin{cases} (0, rx) \cap (0, 1) & \text{if } x \in (0, c), \\ (1 - r(1 - x), 1) \cap (0, 1) & \text{if } x \in (c, 1). \end{cases} \quad (3)$$

The region  $N(c, r, c)$  is arbitrarily taken to be one of  $(0, rx) \cap (0, 1)$  or  $(1 - r(1 - x), 1) \cap (0, 1)$ . Moreover,  $N(0, r, c) := \{0\}$  and  $N(1, r, c) := \{1\}$  for all  $r \geq 1$  and  $c \in [0, 1]$ ; For  $X_i \stackrel{iid}{\sim} \mathcal{U}(y_1, y_2)$ , the special cases in the construction of  $N(\cdot, r, c) - X$  falls at  $c$  or the end points of  $(y_1, y_2)$  — occur with probability zero. Moreover, the region  $N(x, r, c)$  is an interval a.s.

The  $\Gamma_1$ -region,  $\Gamma_1(\mathcal{X}_n, r, c)$ , depends on  $X_{(1)}$ ,  $X_{(n)}$ ,  $r$ , and  $c$ . If  $\Gamma_1(\mathcal{X}_n, r, c) \neq \emptyset$ , then we have  $\Gamma_1(\mathcal{X}_n, r, c) = (\delta_1, \delta_2)$  where at least one end points  $\delta_1, \delta_2$  is a function of  $X_{(1)}$  and  $X_{(n)}$ . For  $\mathcal{U}(0, 1)$  data, given  $X_{(1)} = x_1$  and  $X_{(n)} = x_n$ , the probability of  $p_n(\mathcal{U}, r, c)$  is  $(1 - (\delta_2 - \delta_1)/(x_n - x_1))^{(n-2)}$  provided that  $\Gamma_1(\mathcal{X}_n, r, c) \neq \emptyset$ ; and if  $\Gamma_1(\mathcal{X}_n, r, c) = \emptyset$ , then  $\gamma_{n,2}(\mathcal{U}, r, c) = 2$  holds. Then

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) \neq \emptyset) = \int \int_{\mathcal{S}_1} f_{1n}(x_1, x_n) \left(1 - \frac{\delta_2 - \delta_1}{x_n - x_1}\right)^{(n-2)} dx_n dx_1 \quad (4)$$

where  $\mathcal{S}_1 = \{0 < x_1 < x_n < 1 : x_1, x_n \notin \Gamma_1(\mathcal{X}_n, r, c) \text{ and } \Gamma_1(\mathcal{X}_n, r, c) \neq \emptyset\}$  and  $f_{1n}(x_1, x_n) = n(n-1)[x_n - x_1]^{(n-2)} \mathbf{I}(0 < x_1 < x_n < 1)$ . The integral simplifies to

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) \neq \emptyset) = \int \int_{\mathcal{S}_1} n(n-1)[x_n - x_1 + \delta_1 - \delta_2]^{(n-2)} dx_n dx_1. \quad (5)$$

If  $\Gamma_1(\mathcal{X}_n, r, c) = \emptyset$ , then  $\gamma_{n,2}(\mathcal{U}, r, c) = 2$ . So

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = \emptyset) = \int \int_{\mathcal{S}_2} f_{1n}(x_1, x_n) dx_n dx_1 \quad (6)$$

where  $\mathcal{S}_2 = \{0 < x_1 < x_n < 1 : \Gamma_1(\mathcal{X}_n, r, c) = \emptyset\}$ .

The probability  $p_n(\mathcal{U}, r, c)$  is the sum of the probabilities in Equations (5) and (6).

### 3.1 The Exact Distribution of the Domination Number of $\mathcal{U}(y_1, y_2)$ -random $\mathcal{D}_{n,2}(2, 1/2)$ -digraphs

For  $r = 2$  and  $c = 1/2$ , we have  $N(x, 2, 1/2) = B(x, r(x))$  where  $r(x) = \min(x, 1 - x)$  for  $x \in (0, 1)$ . Hence proportional-edge PCD based on  $N(x, 2, 1/2)$  is equivalent to the CCCD of Priebe et al. (2001). Moreover,  $\Gamma_1(\mathcal{X}_n, 2, 1/2) = (X_{(n)}/2, (1 + X_{(1)})/2)$ . It has been shown that  $p_n(\mathcal{U}, 2, 1/2) = 4/9 - (16/9)4^{-n}$  (Priebe et al. (2001)). Hence, for  $\mathcal{U}(y_1, y_2)$  data with  $n \geq 1$ , we have

$$\gamma_{n,2}(\mathcal{U}, 2, 1/2) = \begin{cases} 1 & \text{w.p. } 5/9 + (16/9)4^{-n}, \\ 2 & \text{w.p. } 4/9 - (16/9)4^{-n}, \end{cases} \quad (7)$$

where w.p. stands for “with probability”. Then as  $n \rightarrow \infty$ ,  $\gamma_{n,2}(\mathcal{U}, 2, 1/2)$  converges in distribution to  $1 + \text{BER}(4/9)$ . For  $m > 2$ , Priebe et al. (2001) computed the exact distribution of  $\gamma_{n,m}(\mathcal{U}, 2, 1/2)$  also. However, the scale invariance property does not hold for general  $F$ ; that is, for  $X_i \stackrel{iid}{\sim} F$  with support  $\mathcal{S}(F) \subseteq (y_1, y_2)$ , the exact and asymptotic distribution of  $\gamma_{n,2}(F, 2, 1/2)$  depends on  $F$  and  $\mathcal{Y}_2$  (Ceyhan (2008)).

### 3.2 The Exact Distribution of the Domination Number of $\mathcal{U}(y_1, y_2)$ -random $\mathcal{D}_{n,2}(2, c)$ -digraphs

For  $r = 2$ ,  $c \in (0, 1)$ , and  $(y_1, y_2) = (0, 1)$ , the  $\Gamma_1$ -region is  $\Gamma_1(\mathcal{X}_n, 2, c) = (X_{(n)}/2, c] \cup [c, (1 + X_{(1)})/2)$ . Notice that  $(X_{(n)}/2, c]$  or  $[c, (1 + X_{(1)})/2)$  could be empty, but not simultaneously.

**Theorem 3.2.** For  $\mathcal{U}(y_1, y_2)$  data and  $n \geq 1$ , we have  $\gamma_{n,2}(\mathcal{U}, 2, c) \sim 1 + \text{BER}(p_n(\mathcal{U}, 2, c))$  where  $p_n(\mathcal{U}, 2, c) = \nu_{1,n}(c)\mathbf{I}(c \in (0, 1/3] + \nu_{2,n}(c)\mathbf{I}(c \in (1/3, 1/2] + \nu_{3,n}(c)\mathbf{I}(c \in (1/2, 2/3] + \nu_{4,n}(c)\mathbf{I}(c \in (2/3, 1))$  with

$$\begin{aligned} \nu_{1,n}(c) &= \frac{2}{3} \left(c + \frac{1}{2}\right)^n - \frac{8}{9}4^{-n} - \frac{2}{3} \left(\frac{1-c}{2}\right)^n + \frac{1}{9}(1-3c)^n - \frac{2}{9} \left(3c - \frac{1}{2}\right)^n, \\ \nu_{2,n}(c) &= \frac{2}{3} \left(c + \frac{1}{2}\right)^n - \frac{8}{9}4^{-n} - \frac{2}{3} \left(\frac{1-c}{2}\right)^n - \frac{2}{9} \left(\frac{3c-1}{2}\right)^n - \frac{2}{9} \left(3c - \frac{1}{2}\right)^n, \end{aligned}$$

$\nu_{3,n}(c) = \nu_{2,n}(1-c)$ , and  $\nu_{4,n}(c) = \nu_{1,n}(1-c)$ . Furthermore,  $\gamma_{n,2}(\mathcal{U}, 2, 0) = \gamma_{n,2}(\mathcal{U}, 2, 1) = 1$  for all  $n \geq 1$ .

Observe that the parameter  $p_n(\mathcal{U}, 2, c)$  is continuous in  $c \in (0, 1)$  for fixed  $n < \infty$ , but there are jumps (hence discontinuities) in  $p_n(\mathcal{U}, 2, c)$  at  $c \in \{0, 1\}$ . In particular,  $\lim_{c \rightarrow 0} p_n(\mathcal{U}, 2, c) = \lim_{c \rightarrow 1} p_n(\mathcal{U}, 2, c) = \lim_{c \rightarrow 0} \nu_{1,n}(c) = \lim_{c \rightarrow 1} \nu_{4,n}(c) = \frac{1}{9} - \frac{2}{9}(-2)^n - \frac{8}{9}4^{-n}$ , but  $p_n(\mathcal{U}, 2, 0) = p_n(\mathcal{U}, 2, 1) = 0$  for all  $n \geq 1$ . For  $c = 1/2$ , we have  $p_n(\mathcal{U}, 2, c) = 4/9 - (16/9)4^{-n}$ , hence the distribution of  $\gamma_{n,2}(\mathcal{U}, 2, c = 1/2)$  is same as in Equation (7).

In the limit as  $n \rightarrow \infty$ , for  $c \in [0, 1]$ , we have

$$\gamma_{n,2}(\mathcal{U}, 2, c) \sim \begin{cases} 1 + \text{BER}(4/9), & \text{for } c = 1/2, \\ 1, & \text{for } c \neq 1/2. \end{cases}$$

Observe also the interesting behavior of the asymptotic distribution of  $\gamma_{n,2}(\mathcal{U}, 2, c)$  around  $c = 1/2$ . The parameter  $p(\mathcal{U}, 2, c)$  is continuous in  $c \in [0, 1] \setminus \{1/2\}$  (in fact it is unity), but there is a jump (hence discontinuity) in  $p(\mathcal{U}, 2, c)$  at  $c = 1/2$ , since  $p(\mathcal{U}, 2, 1/2) = 4/9$  and  $p(\mathcal{U}, 2, c) = 0$  for  $c \neq 1/2$ . Hence for  $c = 1/2$ , the asymptotic distribution is non-degenerate, and for  $c \neq 1/2$ , the asymptotic distribution is degenerate. That is, for  $c = 1/2 \pm \varepsilon$  with  $\varepsilon > 0$  arbitrarily small, although the exact distribution is non-degenerate, the asymptotic distribution is degenerate.

### 3.3 The Exact Distribution of the Domination Number of $\mathcal{U}(y_1, y_2)$ -random $\mathcal{D}_{n,2}(r, 1/2)$ -digraphs

For  $r \geq 1$ ,  $c = 1/2$ , and  $(y_1, y_2) = (0, 1)$ , the  $\Gamma_1$ -region is  $\Gamma_1(\mathcal{X}_n, r, 1/2) = (X_{(n)}/r, 1/2] \cup [1/2, (r-1+X_{(1)})/r)$  where  $(X_{(n)}/r, 1/2]$  or  $[1/2, (r-1+X_{(1)})/r)$  could be empty, but not simultaneously.

**Theorem 3.3.** *For  $\mathcal{U}(y_1, y_2)$  data with  $n \geq 1$ , we have  $\gamma_{n,2}(\mathcal{U}, r, 1/2) \sim 1 + \text{BER}(p_n(\mathcal{U}, r, 1/2))$  where*

$$p_n(\mathcal{U}, r, 1/2) = \begin{cases} \frac{2r}{(r+1)^2} \left( \left( \frac{2}{r} \right)^{n-1} - \left( \frac{r-1}{r^2} \right)^{n-1} \right) & \text{for } r \geq 2, \\ 1 - \frac{1+r^{2n-1}}{(2r)^{n-1}(r+1)} + \frac{(r-1)^n}{(r+1)^2} \left( 1 - \left( \frac{r-1}{2r} \right)^{n-1} \right) & \text{for } 1 \leq r < 2. \end{cases}$$

Notice that for fixed  $n < \infty$ , the parameter  $p_n(\mathcal{U}, r, 1/2)$  is continuous in  $r \geq 2$ . In particular, for  $r = 2$ , we have  $p_n(\mathcal{U}, 2, 1/2) = 4/9 - (16/9)4^{-n}$ , hence the distribution of  $\gamma_{n,2}(\mathcal{U}, r = 2, 1/2)$  is same as in Equation (7). Furthermore,  $\lim_{r \rightarrow 1} p_n(\mathcal{U}, r, 1/2) = p_n(\mathcal{U}, 1, 1/2) = 1 - 2^{1-n}$  and  $\lim_{r \rightarrow \infty} p_n(\mathcal{U}, r, 1/2) = p_n(\mathcal{U}, \infty, 1/2) = 0$ .

In the limit, as  $n \rightarrow \infty$ , we have

$$\gamma_{n,2}(\mathcal{U}, r, 1/2) \sim \begin{cases} 1 & \text{for } r > 2, \\ 1 + \text{BER}(4/9) & \text{for } r = 2, \\ 2 & \text{for } 1 \leq r < 2. \end{cases}$$

Observe the interesting behavior of the asymptotic distribution of  $\gamma_{n,2}(\mathcal{U}, r, 1/2)$  around  $r = 2$ . The parameter  $p(\mathcal{U}, r, 1/2)$  is continuous (in fact piecewise constant) for  $r \in [1, \infty) \setminus \{2\}$ . Hence for  $r \neq 2$ , the asymptotic distribution is degenerate, as  $p(\mathcal{U}, r, 1/2) = 1$  for  $r > 2$  and  $p(\mathcal{U}, r, 1/2) = 2$  w.p. 1 for  $r < 2$ . That is, for  $r = 2 \pm \varepsilon$  with  $\varepsilon > 0$  arbitrarily small, although the exact distribution is non-degenerate, the asymptotic distribution is degenerate.

### 3.4 The Distribution of the Domination Number of $\mathcal{U}(y_1, y_2)$ -random $\mathcal{D}_{n,2}(r, c)$ -digraphs

For  $r \geq 1$  and  $c \in (0, 1)$ , the  $\Gamma_1$ -region is  $\Gamma_1(\mathcal{X}_n, r, c) = (X_{(n)}/r, c] \cup [c, (r-1+X_{(1)})/r)$  where  $(X_{(n)}/r, c]$  or  $[c, (r-1+X_{(1)})/r)$  could be empty, but not simultaneously.

**Theorem 3.4. Main Result 1:** *For  $\mathcal{U}(y_1, y_2)$  data with  $n \geq 1$ ,  $r \geq 1$ , and  $c \in ((3 - \sqrt{5})/2, 1/2)$ , we have  $\gamma_{n,2}(\mathcal{U}, r, c) \sim 1 + \text{BER}(p_n(\mathcal{U}, r, c))$  where  $p_n(\mathcal{U}, r, c) = \pi_{1,n}(r, c) \mathbf{I}(r \geq 1/c) + \pi_{2,n}(r, c) \mathbf{I}(1/(1-c) \leq r < 1/c) + \pi_{3,n}(r, c) \mathbf{I}((1-c)/c \leq r < 1/(1-c)) + \pi_{4,n}(r, c) \mathbf{I}(1 \leq r < (1-c)/c)$  with*

$$\pi_{1,n}(r, c) = \frac{2r}{(r+1)^2} \left( \left( \frac{2}{r} \right)^{n-1} - \left( \frac{r-1}{r^2} \right)^{n-1} \right),$$

$$\pi_{2,n}(r, c) = \frac{1}{(r+1)r^{n-1}} \left[ (1+cr)^n - (1-c)^n - \frac{1}{r+1} (cr^2 - r + cr + 1)^n - \frac{(r-1)^{n-1}}{r+1} \left( \frac{1}{r^{n-2}} + (cr-1+c)^n \right) \right],$$

$$\pi_{3,n}(r, c) = 1 + \frac{(r-1)^{n-1}}{(r+1)^2} \left[ (r-1) - \frac{1}{r^{n-1}} ((cr-1+c)^n + (r-cr-c)^n) \right] - \frac{1}{r+1} [c^n + (1-c)^n] \left( r^n + \frac{1}{r^{n-1}} \right),$$

and

$$\begin{aligned} \pi_{4,n}(r, c) = 1 + \frac{(r-1)^{n-1}}{(r+1)^2} (1-cr-c)^n \left( r^2 - \left( \frac{-1}{r} \right)^{n-1} (r^2-1) \right) + \frac{(r-1)^n}{(r+1)^2} \left( 1-r \left( \frac{r-cr-c}{r} \right)^n \right) \\ - \frac{1}{r+1} [c^n + (1-c)^n] \left( r^n - \frac{1}{r^{n-1}} \right). \end{aligned}$$

And for  $c \in (0, (3-\sqrt{5})/2]$ , we have  $p_n(\mathcal{U}, r, c) = \vartheta_{1,n}(r, c) \mathbf{I}(r \geq 1/c) + \vartheta_{2,n}(r, c) \mathbf{I}((1-c)/c \leq r < 1/c) + \vartheta_{3,n}(r, c) \mathbf{I}(1/(1-c) \leq r < (1-c)/c) + \vartheta_{4,n}(r, c) \mathbf{I}(1 \leq r < 1/(1-c))$  where  $\vartheta_{1,n}(r, c) = \pi_{1,n}(r, c)$ ,  $\vartheta_{2,n}(r, c) = \pi_{2,n}(r, c)$ ,  $\vartheta_{3,n}(r, c) = \pi_{3,n}(r, c)$ , and  $\vartheta_{4,n}(r, c) = \pi_{4,n}(r, c)$ , and

$$\begin{aligned} \vartheta_{3,n}(r, c) = \frac{r}{(r+1)^2} \left[ (r-1)^{n-1} (1-cr-c)^n \left( r + (r^2-1) \left( \frac{-1}{r} \right)^n \right) - \left( \frac{r-1}{r^2} \right)^{n-1} - \left( \frac{cr^2-c+cr+1}{r} \right)^n + \right. \\ \left. \frac{r+1}{r^n} [(1+cr)^n - (1-c)^n] \right]. \end{aligned}$$

Furthermore, we have  $\gamma_{n,2}(\mathcal{U}, r, 0) = \gamma_{n,2}(\mathcal{U}, r, 1) = 1$  for all  $n \geq 1$ .

Some remarks are in order for the Main Result 1. The partitioning of  $c \in (0, 1/2)$  as  $c \in (0, (3-\sqrt{5})/2]$  and  $c \in ((3-\sqrt{5})/2, 1/2)$  is due to the relative positions of  $1/(1-c)$  and  $(1-c)/c$ . For  $c \in ((3-\sqrt{5})/2, 1/2)$ , we have  $1/(1-c) > (1-c)/c$  and for  $c \in (0, (3-\sqrt{5})/2)$ , we have  $1/(1-c) < (1-c)/c$ . At  $c = (3-\sqrt{5})/2$ ,  $1/(1-c) = (1-c)/c = (\sqrt{5}+1)/2$  and only  $\pi_{1,n}(r, (3-\sqrt{5})/2) = \vartheta_{1,n}(r, (3-\sqrt{5})/2)$ ,  $\pi_{2,n}(r, (3-\sqrt{5})/2) = \vartheta_{2,n}(r, (3-\sqrt{5})/2)$ , and  $\pi_{4,n}(r, (3-\sqrt{5})/2) = \vartheta_{4,n}(r, (3-\sqrt{5})/2)$  terms survive. Also, notice the  $(-1)^n$  terms in  $\pi_{4,n}(r, c)$  and  $\vartheta_{3,n}(r, c)$  which might suggest fluctuations of these probabilities as  $n$  changes (increases). However, as  $n$  increases,  $\pi_{4,n}(r, c)$  strictly increases towards 1 (see Figure 1), and  $\vartheta_{3,n}(r, c)$  decreases (strictly decreases for  $n \geq 3$ ) towards 0 (see Figure 2).

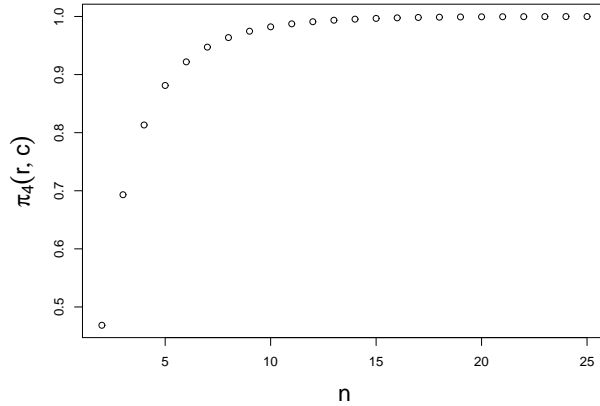


Figure 1: The probability  $\pi_{4,n}(r, c)$  in Main Result 1 with  $r = 1.2$  and  $c = 0.4$  for  $n = 2, 3, \dots, 25$ .

*Remark 3.5.* By symmetry, in Theorem 3.4, for  $c \in (1/2, (\sqrt{5}-1)/2)$ , we have  $p_n(\mathcal{U}, r, c) = \pi_{1,n}(r, 1-c) \mathbf{I}(r \geq 1/(1-c)) + \pi_{2,n}(r, 1-c) \mathbf{I}(1/c \leq r < 1/(1-c)) + \pi_{3,n}(r, 1-c) \mathbf{I}(c/(1-c) \leq r < 1/c) + \pi_{4,n}(r, 1-c) \mathbf{I}(1 \leq r < c/(1-c))$  and for  $c \in [(\sqrt{5}-1)/2, 1)$ ,  $p_n(\mathcal{U}, r, c) = \vartheta_{1,n}(r, 1-c) \mathbf{I}(r \geq 1/(1-c)) + \vartheta_{2,n}(r, 1-c) \mathbf{I}(c/(1-c) \leq r < 1/(1-c)) + \vartheta_{3,n}(r, 1-c) \mathbf{I}(1/c \leq r < c/(1-c)) + \vartheta_{4,n}(r, 1-c) \mathbf{I}(1 \leq r < 1/c)$ .  $\square$

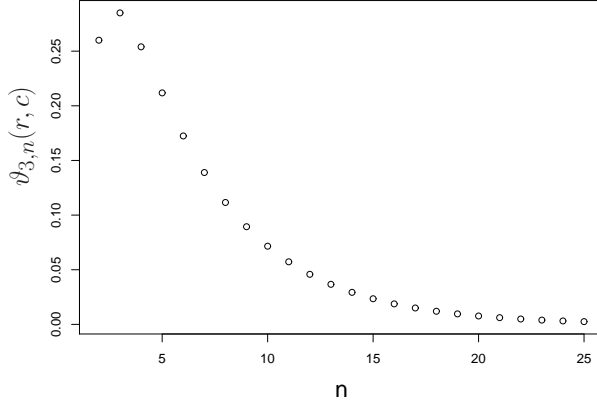


Figure 2: The probability  $\vartheta_{3,n}(r, c)$  in Main Result 1 with  $r = 2$  and  $c = 0.3$  for  $n = 2, 3, \dots, 25$ .

Observe that  $\lim_{r \rightarrow 1} p_n(\mathcal{U}, r, c) = \lim_{r \rightarrow 1} \pi_{4,n}(r, c) = 1$  as expected. For fixed  $1 < n < \infty$ , the probability  $p_n(\mathcal{U}, r, c)$  is continuous in  $(r, c) \in \{(r, c) \in \mathbb{R}^2 : r \geq 1, 0 < c < 1\}$ . In particular, for  $c \in ((3 - \sqrt{5})/2, 1/2)$ , as  $(r, c) \rightarrow (2, 1/2)$  in  $\{(r, c) \in \mathbb{R}^2 : r \geq 1/c\}$ ,  $p_n(\mathcal{U}, r, c) = \pi_{1,n}(r, c) \rightarrow 4/9 - (16/9)4^{-n}$ ; as  $(r, c) \rightarrow (2, 1/2)$  in  $\{(r, c) \in \mathbb{R}^2 : 1/(1-c) \leq r < 1/c\}$ ,  $p_n(\mathcal{U}, r, c) = \pi_{2,n}(r, c) \rightarrow 4/9 - (16/9)4^{-n}$ ; and as  $(r, c) \rightarrow (2, 1/2)$  in  $\{(r, c) \in \mathbb{R}^2 : (1-c)/c \leq r < 1/(1-c)\}$ ,  $p_n(\mathcal{U}, r, c) = \pi_{3,n}(r, c) \rightarrow 4/9 - (16/9)4^{-n}$ . The limit  $(r, c) \rightarrow (2, 1/2)$  is not possible for  $\{(r, c) \in \mathbb{R}^2 : 1 \leq r < (1-c)/c\}$ . For  $c \in (0, (3 - \sqrt{5})/2]$ ,  $(r, c) \rightarrow (2, 1/2)$  can not occur either. And for  $(r, c) = (2, 1/2)$ , the distribution of  $\gamma_{n,2}(\mathcal{U}, r, c)$  is  $1 + \text{BER}(p_n(\mathcal{U}, 2, 1/2))$ , where  $p_n(\mathcal{U}, 2, 1/2) = 4/9 - (16/9)4^{-n}$  as in Equation (7). Therefore for fixed  $1 < n < \infty$ , as  $(r, c) \rightarrow (2, 1/2)$  in  $S = \{(r, c) \in \mathbb{R}^2 : r \geq 1, 0 < c < 1/2\}$ , we have  $p_n(\mathcal{U}, r, c) \rightarrow 4/9 - (16/9)4^{-n}$ . Hence as  $(r, c) \rightarrow (2, 1/2)$  in  $S$ ,  $\gamma_{n,2}(\mathcal{U}, r, c)$  converges in distribution to  $\gamma_{n,2}(\mathcal{U}, 2, 1/2)$ . However,  $p_n(\mathcal{U}, r, c)$  has jumps (hence discontinuities) at  $c \in \{0, 1\}$ . As  $c \rightarrow 0^+$  (which implies we should consider  $c \in (0, (3 - \sqrt{5})/2]$ ,  $1/c \rightarrow \infty$ ,  $(1-c)/c \rightarrow \infty$ , and  $1/(1-c) \rightarrow 1^+$ . Hence  $\lim_{c \rightarrow 0^+} \vartheta_{1,n}(r, c) = \vartheta_{1,n}(\infty, 0) = 0$ ,  $\lim_{c \rightarrow 0^+} \vartheta_{2,n}(r, c) = \vartheta_{2,n}(\infty, 0) = 0$ . Moreover,  $\lim_{c \rightarrow 0^+} \vartheta_{3,n}(r, c) = \frac{r(r-1)^{n-1}}{(r+1)^2} [r + (-r)^{1-n}(1-r) - r^{2-2n}]$ ;  $\lim_{c \rightarrow 0^+} \vartheta_{4,n}(r, c) = 1 + \frac{1}{r+1} [(r-1)^n(1-(-r)^{1-n}) - r^n - r^{1-n}]$ . But  $p_n(\mathcal{U}, r, 0) = 0$  for all  $r \geq 1$ . Similar results can be obtained as  $c \rightarrow 1^-$ . Observe also that  $\lim_{r \rightarrow 1} p_n(\mathcal{U}, r, c) = \lim_{r \rightarrow 1} \pi_{4,n}(r, c) = 1$ .

**Theorem 3.6. Main Result 2:** Let  $D_{n,2}(r, c)$  be based on  $\mathcal{U}(y_1, y_2)$  data with  $c \in (0, 1)$  and  $\tau = \max(c, 1-c)$ . Then the domination number  $\gamma_{n,2}(\mathcal{U}, r, c)$  of the PCD has the following asymptotic distribution. As  $n \rightarrow \infty$ ,

$$\gamma_{n,2}(\mathcal{U}, r, c) \sim \begin{cases} 1 + \text{BER}(r/(r+1)), & \text{for } r = 1/\tau, \\ 1, & \text{for } r > 1/\tau, \\ 2, & \text{for } 1 \leq r < 1/\tau. \end{cases} \quad (8)$$

Notice the interesting behavior of the asymptotic distribution of  $\gamma_{n,2}(\mathcal{U}, r, c)$  around  $r = 1/\tau$  for any given  $c \in (0, 1)$ . The asymptotic distribution is non-degenerate only for  $r = 1/\tau$ . For  $r > 1/\tau$ ,  $\lim_{n \rightarrow \infty} \gamma_{n,2}(\mathcal{U}, r, c) = 1$  w.p. 1, and for  $1 \leq r < 1/\tau$ ,  $\lim_{n \rightarrow \infty} \gamma_{n,2}(\mathcal{U}, r, 1/2) = 2$  w.p. 1. The critical value  $r = 1/\tau$  corresponds to  $c = (r-1)/r$ , if  $c \in (0, 1/2)$  (i.e.,  $\tau = 1-c$ ) and  $c = 1/r$ , if  $c \in (1/2, 1)$  (i.e.,  $\tau = c$ ) and these are only possible for  $r \in (1, 2)$ . That is, for  $r = (1/\tau) \pm \varepsilon$  for  $\varepsilon$  arbitrarily small, although the exact distribution is non-degenerate, the asymptotic distribution is degenerate. The parameter  $p(\mathcal{U}, r, c)$  is continuous in  $r$  and  $c$  for  $(r, c) \in S \setminus \{1/\tau, c\}$  and there is a jump (hence discontinuity) in the probability  $p(\mathcal{U}, r, c)$  at  $r = 1/\tau$ , since  $p(\mathcal{U}, 1/\tau, c) = 1/(1+\tau) = r/(r+1)$ . Therefore, given a centrality parameter  $c \in (0, 1)$ , we can choose the expansion parameter  $r$  for which the asymptotic distribution is non-degenerate, and vice versa.

There is yet another interesting behavior of the asymptotic distribution around  $(r, c) = (2, 1/2)$ . The parameter  $p(\mathcal{U}, r, c)$  has jumps at  $c = 1/r$  and  $(r-1)/r$  for  $r \in [1, 2]$  with  $p(\mathcal{U}, r, 1/r) = p(\mathcal{U}, r, (r-1)/r) = r/(r+1)$ . That is, for fixed  $(r, c) \in S$ ,  $\lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, (r-1)/r) = \lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, 1/r) = r/(r+1)$ . Letting

$(r, c) \rightarrow (2, 1/2)$  (i.e.,  $r \rightarrow 2$ ) we get  $p(\mathcal{U}, r, (r-1)/r) \rightarrow 2/3$  and  $p(\mathcal{U}, r, 1/r) \rightarrow 2/3$ , but  $p(\mathcal{U}, 2, 1/2) = 4/9$ . Hence for  $r \in [1, 2)$  the distributions of  $\gamma_{n,2}(\mathcal{U}, r, (r-1)/r)$  and  $\gamma_{n,2}(\mathcal{U}, r, 1/r)$  are identical and both converge to  $1 + \text{BER}(r/(r+1))$ , but the distribution of  $\gamma_{n,2}(\mathcal{U}, 2, 1/2)$  converges to  $1 + \text{BER}(4/9)$  as  $n \rightarrow \infty$ . In other words,  $p(\mathcal{U}, r, (r-1)/r) = p(\mathcal{U}, r, 1/r)$  has another jump at  $r = 2$  (which corresponds to  $(r, c) = (2, 1/2)$ ). This interesting behavior might be due to the symmetry around  $c = 1/2$ . Because for  $c \in (0, 1/2)$ , with  $r = 1/(1-c)$ , for sufficiently large  $n$ , a point  $X_i$  in  $(c, 1)$  can dominate all the points in  $\mathcal{X}_n$  (implying  $\gamma_{n,2}(\mathcal{U}, r, (r-1)/r) = 1$ ), but no point in  $(0, c)$  can dominate all points a.s. Likewise, for  $c \in (1/2, 1)$  with  $r = 1/c$ , for sufficiently large  $n$ , a point  $X_i$  in  $(0, c)$  can dominate all the points in  $\mathcal{X}_n$  (implying  $\gamma_{n,2}(\mathcal{U}, r, 1/r) = 1$ ), but no point in  $(c, 1)$  can dominate all points a.s. However, for  $c = 1/2$  and  $r = 2$ , for sufficiently large  $n$ , points to the left or right of  $c$  can dominate all other points in  $\mathcal{X}_n$ .

## 4 The Distribution of the Domination Number for $\mathcal{F}(\mathbb{R})$ -random $\mathcal{D}_{n,2}(r, c)$ -digraphs

Let  $\mathcal{F}(y_1, y_2)$  be a family of continuous distributions with support in  $\mathcal{S}_F \subseteq (y_1, y_2)$ . Consider a distribution function  $F \in \mathcal{F}(y_1, y_2)$ . For simplicity, assume  $y_1 = 0$  and  $y_2 = 1$ . Let  $\mathcal{X}_n$  be a random sample from  $F$ ,  $\Gamma_1(\mathcal{X}_n, r, c) = (\delta_1, \delta_2)$ ,  $p_n(F, r, c) := P(\gamma_{n,2}(F, r, c) = 2)$ , and  $p(F, r, c) := \lim_{n \rightarrow \infty} P(\gamma_{n,2}(F, r, c) = 2)$ . The exact (i.e., finite sample) and asymptotic distributions of  $\gamma_{n,2}(F, r, c)$  are  $1 + \text{BER}(p_n(F, r, c))$  and  $1 + \text{BER}(p(F, r, c))$ , respectively. That is, for finite  $n > 1$ ,  $r \in [1, \infty)$ , and  $c \in (0, 1)$ , we have

$$\gamma_{n,2}(F, r, c) = \begin{cases} 1 & \text{w.p. } 1 - p_n(F, r, c), \\ 2 & \text{w.p. } p_n(F, r, c). \end{cases} \quad (9)$$

Moreover,  $\gamma_{1,2}(F, r, c) = 1$  for all  $r \geq 1$  and  $c \in [0, 1]$ ,  $\gamma_{n,2}(F, r, 0) = \gamma_{1,2}(F, r, 1) = 1$  for all  $n \geq 1$  and  $r \geq 1$ ,  $\gamma_{n,2}(F, \infty, c) = 1$  for all  $n \geq 1$  and  $c \in [0, 1]$ , and  $\gamma_{n,2}(F, 1, c) = k_4$  for all  $n \geq 1$  and  $c \in (0, 1)$  where  $k_4$  is as in Theorem 2.5 with  $m = 2$ . The asymptotic distribution is similar with  $p_n(F, r, c)$  being replaced by  $p(F, r, c)$ . The special cases are similar in the asymptotics with the exception that  $p(F, 1, c) = 1$  for all  $c \in (0, 1)$ . The finite sample mean and variance of  $\gamma_{n,2}(F, r, c)$  are given by  $1 + p_n(F, r, c)$  and  $p_n(F, r, c)(1 - p_n(F, r, c))$ , respectively; and the asymptotic mean and variance of  $\gamma_{n,2}(F, r, c)$  are given by  $1 + p(F, r, c)$  and  $p(F, r, c)(1 - p(F, r, c))$ , respectively.

Given  $X_{(1)} = x_1$  and  $X_{(n)} = x_n$ , the probability of  $\gamma_{n,2}(F, r, c) = 2$  (i.e.,  $p_n(F, r, c)$ ) is  $(1 - [F(\delta_2) - F(\delta_1)]/[F(x_n) - F(x_1)])^{(n-2)}$  provided that  $\Gamma_1(\mathcal{X}_n, r, c) = (\delta_1, \delta_2) \neq \emptyset$ ; if  $\Gamma_1(\mathcal{X}_n, r, c) = \emptyset$ , then  $\gamma_{n,2}(F, r, c) = 2$ . Then

$$P(\gamma_{n,2}(F, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) \neq \emptyset) = \int \int_{\mathcal{S}_1} f_{1n}(x_1, x_n) \left(1 - \frac{F(\delta_2) - F(\delta_1)}{F(x_n) - F(x_1)}\right)^{(n-2)} dx_n dx_1 \quad (10)$$

where  $\mathcal{S}_1 = \{0 < x_1 < x_n < 1 : (x_1, x_n) \notin \Gamma_1(\mathcal{X}_n, r, c), \Gamma_1(\mathcal{X}_n, r, c) \neq \emptyset\}$  and  $f_{1n}(x_1, x_n) = n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1)]^{(n-2)}\mathbf{I}(0 < x_1 < x_n < 1)$  which is the joint pdf of  $X_{(1)}, X_{(n)}$ . The integral in (10) becomes

$$P(\gamma_{n,2}(F, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) \neq \emptyset) = \int \int_{\mathcal{S}_1} H(x_1, x_n) dx_n dx_1, \quad (11)$$

where

$$H(x_1, x_n) := n(n-1)f(x_1)f(x_n)[F(x_n) + F(\delta_1) - (F(\delta_2) + F(x_1))]^{n-2}. \quad (12)$$

If  $\Gamma_1(\mathcal{X}_n, r, c) = \emptyset$ , then  $\gamma_{n,2}(F, r, c) = 2$ . So

$$P(\gamma_{n,2}(F, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = \emptyset) = \int \int_{\mathcal{S}_2} f_{1n}(x_1, x_n) dx_n dx_1 \quad (13)$$

where  $\mathcal{S}_2 = \{0 < x_1 < x_n < 1 : \Gamma_1(\mathcal{X}_n, r, c) = \emptyset\}$ .

The probability  $p_n(F, r, c)$  is the sum of the probabilities in Equations (11) and (13).

For  $\mathcal{Y}_2 = \{y_1, y_2\} \subset \mathbb{R}$  with  $-\infty < y_1 < y_2 < \infty$ , a quick investigation shows that the  $\Gamma_1$ -region is  $\Gamma_1(\mathcal{X}_n, r, c) = (y_1 + (X_{(n)} - y_1)/r, M_c] \cup [M_c, y_2 - (y_2 - X_{(1)})/r)$ . Notice that for a given  $c \in [0, 1]$ , the corresponding  $M_c \in [y_1, y_2]$  is  $M_c = y_1 + c(y_2 - y_1)$ .

Let  $F$  be a continuous distribution with support  $\mathcal{S}(F) \subseteq (0, 1)$ . The simplest of such distributions is  $\mathcal{U}(0, 1)$ , which yields the simplest exact distribution for  $\gamma_{n,2}(F, r, c)$  with  $(r, c) = (1, 1/2)$ . If  $X \sim F$ , then by probability integral transform,  $F(X) \sim \mathcal{U}(0, 1)$ . So for any continuous  $F$ , we can construct a proximity map depending on  $F$  for which the distribution of the domination number of the associated digraph has the same distribution as that of  $\gamma_{n,2}(\mathcal{U}, r, c)$ .

**Proposition 4.1.** *Let  $X_i \stackrel{iid}{\sim} F$  which is an absolutely continuous distribution with support  $\mathcal{S}(F) = (0, 1)$  and  $\mathcal{X}_n = \{X_1, X_2, \dots, X_n\}$ . Define the proximity map  $N_F(x, r, c) := F^{-1}(N(F(x), r, c))$ . Then the domination number of the digraph based on  $N_F$ ,  $\mathcal{X}_n$ , and  $\mathcal{Y}_2 = \{0, 1\}$  has the same distribution as  $\gamma_{n,2}(\mathcal{U}, r, c)$ .*

**Proof:** Let  $U_i := F(X_i)$  for  $i = 1, 2, \dots, n$  and  $\mathcal{U}_n := \{U_1, U_2, \dots, U_n\}$ . Hence, by probability integral transform,  $U_i \stackrel{iid}{\sim} \mathcal{U}(0, 1)$ . Let  $U_{(i)}$  be the  $i^{th}$  order statistic of  $\mathcal{U}_n$  for  $i = 1, 2, \dots, n$ . Furthermore, an absolutely continuous  $F$  preserves order; that is, for  $x \leq y$ , we have  $F(x) \leq F(y)$ . So the image of  $N_F(x, r, c)$  under  $F$  is  $F(N_F(x, r, c)) = N(F(x), r, c)$  for (almost) all  $x \in (0, 1)$ . Then  $F(N_F(X_i, r, c)) = N(F(X_i), r, c) = N(U_i, r, c)$  for  $i = 1, 2, \dots, n$ . Since  $U_i \stackrel{iid}{\sim} \mathcal{U}(0, 1)$ , the distribution of the domination number of the digraph based on  $N(\cdot, r, c)$ ,  $\mathcal{U}_n$ , and  $\{0, 1\}$  is given in Theorem 3.4. Observe that for any  $j$ ,  $X_j \in N_F(X_i, r, c)$  iff  $X_j \in F^{-1}(N(F(X_i), r, c))$  iff  $F(X_j) \in N(F(X_i), r, c)$  iff  $U_j \in N(U_i, r, c)$  for  $i = 1, 2, \dots, n$ . Hence  $P(\mathcal{X}_n \subset N_F(X_i, r, c)) = P(\mathcal{U}_n \subset N(U_i, r, c))$  for all  $i = 1, 2, \dots, n$ . Therefore,  $\mathcal{X}_n \cap \Gamma_1(\mathcal{X}_n, N_F(r, c)) = \emptyset$  iff  $\mathcal{U}_n \cap \Gamma_1(\mathcal{U}_n, r, c) = \emptyset$ . Hence the desired result follows. ■

There is also a stochastic ordering between  $\gamma_{n,2}(F, r, c)$  and  $\gamma_{n,2}(\mathcal{U}, r, c)$  provided that  $F$  satisfies some regularity conditions.

**Proposition 4.2.** *Let  $\mathcal{X}_n = \{X_1, X_2, \dots, X_n\}$  be a random sample from an absolutely continuous distribution  $F$  with  $\mathcal{S}(F) \subseteq (0, 1)$ . If*

$$F(X_{(n)})/r < F(X_{(n)})/r \text{ and } F(X_{(1)}) < rF((X_{(1)} + r - 1)/r) + 1 - r \text{ hold a.s.,} \quad (14)$$

*then  $\gamma_{n,2}(F, r, c) <^{ST} \gamma_{n,2}(\mathcal{U}, r, F(c))$  where  $<^{ST}$  stands for “stochastically smaller than”. If  $<$ ’s in (14) are replaced with  $>$ ’s, then  $\gamma_{n,2}(\mathcal{U}, r, F(c)) <^{ST} \gamma_{n,2}(F, r, c)$ . If  $<$ ’s in expression (14) are replaced with  $=$ ’s, then  $\gamma_{n,2}(F, r, c) \stackrel{d}{=} \gamma_{n,2}(\mathcal{U}, r, F(c))$  where  $\stackrel{d}{=}$  stands for equality in distribution.*

**Proof:** Let  $U_i$  and  $U_{(i)}$  be as in proof of Proposition 4.1. Then the parameter  $c$  for  $N(\cdot, r, c)$  with  $\mathcal{X}_n$  in  $(0, 1)$  corresponds to  $F(c)$  for  $\mathcal{U}_n$ . Then the  $\Gamma_1$ -region for  $\mathcal{U}_n$  based on  $N(\cdot, r, F(c))$  is  $\Gamma_1(\mathcal{U}_n, r, F(c)) = (U_{(n)}/r, F(c)] \cup [F(c), (U_{(1)} + r - 1)/r)$ ; likewise,  $\Gamma_1(\mathcal{X}_n, r, c) = (X_{(n)}/r, M_c] \cup [M_c, (X_{(1)} + r - 1)/r)$ . But the conditions in (14) imply that  $\Gamma_1(\mathcal{U}_n, r, F(c)) \subsetneq F(\Gamma_1(\mathcal{X}_n, r, c))$ , since such an  $F$  preserves order. So  $\mathcal{U}_n \cap F(\Gamma_1(\mathcal{X}_n, r, c)) = \emptyset$  implies that  $\mathcal{U}_n \cap \Gamma_1(\mathcal{U}_n, r, F(c)) = \emptyset$  and  $\mathcal{U}_n \cap F(\Gamma_1(\mathcal{X}_n, r, c)) = \emptyset$  iff  $\mathcal{X}_n \cap \Gamma_1(\mathcal{X}_n, r, F(c)) = \emptyset$ . Hence

$$p_n(F, r, c) = P(\mathcal{X}_n \cap \Gamma_1(\mathcal{X}_n, r, c) = \emptyset) < P(\mathcal{U}_n \cap \Gamma_1(\mathcal{U}_n, r, F(c)) = \emptyset) = p_n(\mathcal{U}, r, F(c)).$$

Then  $\gamma_{n,2}(F, r, c) <^{ST} \gamma_{n,2}(\mathcal{U}, r, F(c))$  follows. The other cases follow similarly. ■

*Remark 4.3.* We can also find the exact distribution of  $\gamma_{n,2}(F, r, c)$  for  $F$  whose pdf is piecewise constant with support in  $(0, 1)$  as in Ceyhan (2008). Note that the simplest of such distributions is the uniform distribution  $\mathcal{U}(0, 1)$ . The exact distribution of  $\gamma_{n,2}(F, r, c)$  for (piecewise) polynomial  $f(x)$  with at least one piece is of degree 1 or higher and support in  $(0, 1)$  can be obtained using the multinomial expansion of the term  $(\cdot)^{n-2}$  in Equation (12) with careful bookkeeping. However, the resulting expression for  $p_n(F, r, c)$  is extremely lengthy and not that informative (see Ceyhan (2008)).

For fixed  $n$ , one can obtain  $p_n(F, r, c)$  for  $F$  (omitted for the sake of brevity) by numerical integration of the below expression.

$$p_n(F, r, c) = P(\gamma_{n,2}(F, r, c) = 2) = \int \int_{\mathcal{S}(F) \setminus (\delta_1, \delta_2)} H(x_1, x_n) dx_n dx_1,$$

where  $H(x_1, x_n)$  is given in Equation (12).  $\square$

Recall the  $\mathcal{F}(\mathbb{R}^d)$ -random  $\mathcal{D}_{n,m}(r, c)$ -digraphs. We call the digraph which obtains in the special case of  $\mathcal{Y}_2 = \{y_1, y_2\}$  and support of  $F_X$  in  $(y_1, y_2)$ ,  $\mathcal{F}(y_1, y_2)$ -random  $\mathcal{D}_{n,2}(r, c)$ -digraph. Below, we provide asymptotic results pertaining to the distribution of such digraphs.

#### 4.1 The Asymptotic Distribution of the Domination Number of $\mathcal{F}(y_1, y_2)$ -random $\mathcal{D}_{n,2}(r, c)$ -digraphs

Although the exact distribution of  $\gamma_{n,2}(F, r, c)$  is not analytically available in a simple closed form for  $F$  whose density is not piecewise constant, the asymptotic distribution of  $\gamma_{n,2}(F, r, c)$  is available for larger families of distributions. First, we present the asymptotic distribution of  $\gamma_{n,2}(F, r, c)$  for  $\mathcal{D}_{n,2}(r, c)$ -digraphs with  $\mathcal{Y}_2 = \{y_1, y_2\} \subset \mathbb{R}$  with  $y_1 < y_2$  for general  $F$  with support  $\mathcal{S}(F) \subseteq (y_1, y_2)$ . Then we will extend this to the case with  $\mathcal{Y}_m \subset \mathbb{R}$  with  $m > 2$ .

Let  $c \in (0, 1/2)$  and  $r \in (1, 2)$ . Then for  $c = (r - 1)/r$ , i.e.,  $M_c = y_1 + (r - 1)(y_2 - y_1)/r$ , we define the family of distributions

$$\mathcal{F}_1(y_1, y_2) := \left\{ F : (y_1, y_1 + \varepsilon) \cup (M_c, M_c + \varepsilon) \subseteq \mathcal{S}(F) \subseteq (y_1, y_2) \text{ for some } \varepsilon \in (0, c) \text{ with } c = (r - 1)/r \right\}.$$

Similarly, let  $c \in (1/2, 1)$  and  $r \in (1, 2)$ . Then for  $c = 1/r$ , i.e.,  $M_c = y_1 + (y_2 - y_1)/r$  with  $r \in (1, 2)$ , we define

$$\mathcal{F}_2(y_1, y_2) := \left\{ F : (y_2 - \varepsilon, y_2) \cup (M_c - \varepsilon, M_c) \subseteq \mathcal{S}(F) \subseteq (y_1, y_2) \text{ for some } \varepsilon \in (0, 1 - c) \text{ with } c = 1/r \right\}.$$

Let the  $k^{\text{th}}$  order right (directed) derivative at  $x$  be defined as  $f^{(k)}(x^+) := \lim_{h \rightarrow 0^+} \frac{f^{(k-1)}(x+h) - f^{(k-1)}(x)}{h}$  for all  $k \geq 1$  and the right limit at  $u$  be defined as  $f(u^+) := \lim_{h \rightarrow 0^+} f(u + h)$ . The left derivatives and limits are defined similarly with  $+$ 's being replaced by  $-$ 's.

**Theorem 4.4. Main Result 3:** Let  $\mathcal{Y}_2 = \{y_1, y_2\} \subset \mathbb{R}$  with  $-\infty < y_1 < y_2 < \infty$ ,  $\mathcal{X}_n = \{X_1, X_2, \dots, X_n\}$  with  $X_i \stackrel{iid}{\sim} F \in \mathcal{F}_1(y_1, y_2)$ , and  $c \in (0, 1/2)$ . Let  $D_{n,2}$  be the  $\mathcal{F}_1(y_1, y_2)$ -random  $\mathcal{D}_{n,2}(r, c)$ -digraph based on  $\mathcal{X}_n$  and  $\mathcal{Y}_2$ .

- (i) Then for  $n > 1$ ,  $r \in (1, \infty)$ , and  $c = (r - 1)/r$  we have  $\gamma_{n,2}(F, r, (r - 1)/r) \sim 1 + \text{BER}(p_n(F, r, (r - 1)/r))$ . Note also that  $\gamma_{1,2}(F, r, (r - 1)/r) = 1$  for all  $r \geq 1$ ; for  $r = 1$ , we have  $\gamma_{n,2}(F, 1, 0) = 1$  for all  $n \geq 1$  and for  $r = \infty$ , we have  $\gamma_{n,2}(F, \infty, 1) = 1$  for all  $n \geq 1$ .
- (ii) Furthermore, suppose  $k \geq 0$  is the smallest integer for which  $F(\cdot)$  has continuous right derivatives up to order  $(k + 1)$  at  $y_1$ ,  $y_1 + (r - 1)(y_2 - y_1)/r$ , and  $f^{(k)}(y_1^+) + r^{-(k+1)} f^{(k)}\left(\left((r - 1)(y_2 - y_1)/r\right)^+\right) \neq 0$  and  $f^{(i)}(y_1^+) = f^{(i)}\left(\left(y_1 + (r - 1)(y_2 - y_1)/r\right)^+\right) = 0$  for all  $i = 0, 1, 2, \dots, (k - 1)$  and suppose also that  $F(\cdot)$  has a continuous left derivative at  $y_2$ . Then for bounded  $f^{(k)}(\cdot)$ ,  $c = (r - 1)/r$ , and  $r \in (1, 2)$ , we have the following limit

$$p(F, r, (r - 1)/r) = \lim_{n \rightarrow \infty} p_n(F, r, (r - 1)/r) = \frac{f^{(k)}(y_1^+)}{f^{(k)}(y_1^+) + r^{-(k+1)} f^{(k)}\left(\left(y_1 + (r - 1)(y_2 - y_1)/r\right)^+\right)}.$$

Note that in Theorem 4.4

- with  $(y_1, y_2) = (0, 1)$ , we have  $p(F, r, (r-1)/r) = \frac{f^{(k)}(0^+)}{f^{(k)}(0^+) + r^{-(k+1)} f^{(k)}((r-1)/r)^+}$ ,
- if  $f^{(k)}(y_1^+) = 0$  and  $f^{(k)}\left((y_1 + (r-1)(y_2 - y_1)/r)^+\right) \neq 0$ , then  $p_n(F, r, (r-1)/r) \rightarrow 0$  as  $n \rightarrow \infty$ , at rate  $O(\kappa_1(f) \cdot n^{-(k+2)/(k+1)})$  where  $\kappa_1(f)$  is a constant depending on  $f$  and
- if  $f^{(k)}(y_1^+) \neq 0$  and  $f^{(k)}\left((y_1 + (r-1)(y_2 - y_1)/r)^+\right) = 0$ , then  $p_n(F, r, (r-1)/r) \rightarrow 1$  as  $n \rightarrow \infty$ , at rate  $O(\kappa_1(f) \cdot n^{-(k+2)/(k+1)})$ .

**Theorem 4.5. Main Result 4:** Let  $\mathcal{Y}_2 = \{y_1, y_2\} \subset \mathbb{R}$  with  $-\infty < y_1 < y_2 < \infty$ ,  $\mathcal{X}_n = \{X_1, X_2, \dots, X_n\}$  with  $X_i \stackrel{iid}{\sim} F \in \mathcal{F}_2(y_1, y_2)$ , and  $c \in (1/2, 1)$ . Let  $D_{n,2}$  be the  $\mathcal{F}_2(y_1, y_2)$ -random  $\mathcal{D}_{n,2}(r, c)$ -digraph based on  $\mathcal{X}_n$  and  $\mathcal{Y}_2$ .

- (i) Then for  $n > 1$ ,  $r \in (1, \infty)$ , and  $c = 1/r$  we have  $\gamma_{n,2}(F, r, 1/r) \sim 1 + \text{BER}(p_n(F, r, 1/r))$ . Note also that  $\gamma_{1,2}(F, r, 1/r) = 1$  for all  $r \geq 1$ ; for  $r = 1$ , we have  $\gamma_{n,2}(F, 1, 1) = 1$  for all  $n \geq 1$  and for  $r = \infty$ , we have  $\gamma_{n,2}(F, \infty, 0) = 1$  for all  $n \geq 1$ .
- (ii) Furthermore, suppose  $\ell \geq 0$  is the smallest integer for which  $F(\cdot)$  has continuous left derivatives up to order  $(\ell + 1)$  at  $y_2$ , and  $y_1 + (y_2 - y_1)/r$ , and  $f^{(\ell)}(y_2^-) + r^{-(\ell+1)} f^{(\ell)}\left((y_1 + (y_2 - y_1)/r)^-\right) \neq 0$  and  $f^{(i)}(y_2^-) = f^{(i)}\left((y_1 + (y_2 - y_1)/r)^-\right) = 0$  for all  $i = 0, 1, 2, \dots, (\ell - 1)$  and suppose also that  $F(\cdot)$  has a continuous right derivative at  $y_1$ . Additionally, for bounded  $f^{(\ell)}(\cdot)$ ,  $c = 1/r$ , and  $r \in (1, 2)$  we have the following limit

$$p(F, r, 1/r) = \lim_{n \rightarrow \infty} p_n(F, r, 1/r) = \frac{f^{(\ell)}(y_2^-)}{f^{(\ell)}(y_2^-) + r^{-(\ell+1)} f^{(\ell)}\left((y_1 + (y_2 - y_1)/r)^-\right)}.$$

Note that in Theorem 4.5

- with  $(y_1, y_2) = (0, 1)$ , we have  $p(F, r, 1/r) = \frac{f^{(\ell)}(1^-)}{f^{(\ell)}(1^-) + r^{-(\ell+1)} f^{(\ell)}((1/r)^-)}$ ,
- if  $f^{(\ell)}(y_2^-) = 0$  and  $f^{(\ell)}\left((y_1 + (y_2 - y_1)/r)^-\right) \neq 0$ , then  $p_n(F, r, 1/r) \rightarrow 0$  as  $n \rightarrow \infty$ , at rate  $O(\kappa_2(f) \cdot n^{-(\ell+2)/(\ell+1)})$  where  $\kappa_2(f)$  is a constant depending on  $f$  and
- if  $f^{(\ell)}(y_2^-) \neq 0$  and  $f^{(\ell)}\left((y_1 + (y_2 - y_1)/r)^-\right) = 0$ , then  $p_n(F, r, 1/r) \rightarrow 1$  as  $n \rightarrow \infty$ , at rate  $O(\kappa_2(f) \cdot n^{-(\ell+2)/(\ell+1)})$ .

*Remark 4.6.* In Theorems 4.4 and 4.5, we assume that  $f^{(k)}(\cdot)$  and  $f^{(\ell)}(\cdot)$  are bounded on  $(y_1, y_2)$ , respectively. If  $f^{(k)}(\cdot)$  is not bounded on  $(y_1, y_2)$  for  $k \geq 0$ , in particular at  $y_1$ , and  $y_1 + (r-1)(y_2 - y_1)/r$ , for example,  $\lim_{x \rightarrow y_1^+} f^{(k)}(x) = \infty$ , then we have

$$p(F, r, (r-1)/r) = \lim_{\delta \rightarrow 0^+} \frac{f^{(k)}(y_1 + \delta)}{[f^{(k)}(y_1 + \delta) + r^{-(k+1)} f^{(k)}((y_1 + (r-1)(y_2 - y_1)/r) + \delta)]}.$$

If  $f^{(\ell)}(\cdot)$  is not bounded on  $(y_1, y_2)$  for  $\ell \geq 0$ , in particular at  $y_1 + (y_2 - y_1)/r$ , and  $y_2$ , for example,  $\lim_{x \rightarrow y_2^-} f^{(\ell)}(x) = \infty$ , then we have

$$p(F, r, 1/r) = \lim_{\delta \rightarrow 0^+} \frac{f^{(\ell)}(y_2 - \delta)}{[f^{(\ell)}(y_2 - \delta) + r^{-(\ell+1)} f^{(\ell)}((y_1 + (y_2 - y_1)/r) - \delta)]}. \quad \square$$

*Remark 4.7.* The rates of convergence in Theorems 4.4 and 4.5 depends on  $f$ . From the proofs of Theorems 4.4 and 4.5, it follows that for sufficiently large  $n$ ,

$$p_n(F, r, (r-1)/r) \approx p(F, r, (r-1)/r) + \frac{\kappa_1(f)}{n^{-(k+2)/(k+1)}} \text{ and } p_n(F, r, 1/r) \approx p(F, r, 1/r) + \frac{\kappa_2(f)}{n^{-(\ell+2)/(\ell+1)}},$$

where  $\kappa_1(f) = \frac{s_1 s_3^{\frac{1}{k+1}} + s_2 \Gamma(\frac{k+2}{k+1})}{(k+1) s_3^{\frac{k+2}{k+1}}}$  with  $\Gamma(x) = \int_0^\infty e^{-t} t^{(x-1)} dt$ ,  $s_1 = \frac{1}{n^{k+1} k!} f^{(k)}(y_1^+)$ ,  $s_2 = \frac{1}{n^{k+1} k!} f^{(k+1)}(y_1^+)$ ,

and  $s_3 = \frac{1}{(k+1)!} p(F, r, (r-1)/r)$ ,  $\kappa_2(f) = \frac{q_1 \Gamma(\frac{\ell+2}{\ell+1}) + q_2 q_3^{\frac{1}{\ell+1}}}{(\ell+1) q_3^{\frac{\ell+2}{\ell+1}}}$ ,  $q_1 = \frac{(-1)^{\ell+1}}{n^{\ell+1} \ell!} f^{(\ell+1)}(y_2^-)$ ,  $q_2 = \frac{(-1)^\ell}{n^{\ell+1} \ell!} f^{(\ell)}(y_2^-)$ , and  $q_3 = \frac{(-1)^{\ell+1}}{(\ell+1)!} p(F, r, 1/r)$  provided the derivatives exist.  $\square$

The conditions of the Theorems 4.4 and 4.5 might seem a bit esoteric. However, most of the well known functions that are scaled and properly transformed to be pdf of some random variable with support in  $(y_1, y_2)$  satisfy the conditions for some  $k$  or  $\ell$ , hence one can compute the corresponding limiting probabilities  $p(F, r, (r-1)/r)$  and  $p(F, r, 1/r)$ .

**Example 4.8.** (a) For example, with  $F = \mathcal{U}(y_1, y_2)$ , in Theorem 4.4, we have  $k = 0$  and  $f(y_1^+) = f((y_1 + (r-1)(y_2 - y_1)/r)^+) = 1/(y_2 - y_1)$ , and in Theorem 4.5, we have  $\ell = 0$  and  $f(y_2^-) = f((y_1 + (y_2 - y_1)/r)^-) = 1/(y_2 - y_1)$ . Then  $\lim_{n \rightarrow \infty} p_n(F, r, (r-1)/r) = \lim_{n \rightarrow \infty} p_n(F, r, 1/r) = r/(r+1)$ , which agrees with the result given in Equation (??).

(b) For  $F$  with pdf  $f(x) = (x+1/2) \mathbf{I}(0 < x < 1)$ , we have  $k = 0$ ,  $f(0^+) = 1/2$ , and  $f((\frac{r-1}{r})^+) = 3/2 - 1/r$  in Theorem 4.4. Then  $p(F, r, (r-1)/r) = \frac{r^2}{r^2 + 3r - 2}$ . As for Theorem 4.5, we have  $\ell = 0$ ,  $f(1^-) = 3/2$  and  $f((\frac{1}{r})^-) = 1/r + 1/2$ . Then  $p(F, r, 1/r) = \frac{3r^2}{3r^2 + r + 2}$ .

(c) For  $F$  with pdf  $f(x) = (\pi/2) |\sin(2\pi x)| \mathbf{I}(0 < x < 1) = (\pi/2)(\sin(2\pi x) \mathbf{I}(0 < x \leq 1/2) - \sin(2\pi x) \mathbf{I}(1/2 < x < 1))$ , we have  $k = 0$ ,  $f(0^+) = 0$ , and  $f((\frac{r-1}{r})^+) = (\pi/2)(\sin(2\pi(r-1)/r))$  in Theorem 4.4. Then  $p(F, r, (r-1)/r) = 0$ . As for Theorem 4.5, we have  $\ell = 0$ ,  $f(1^-) = 0$  and  $f((\frac{1}{r})^-) = -(\pi/2)(\sin(2\pi/r))$ . Then  $p(F, r, 1/r) = 0$ .

(d) For  $F$  with pdf  $f(x) = \frac{\pi r}{4(r-1)} \sin(\pi r x / (r-1)) \mathbf{I}(0 < x \leq (r-1)/r) + g(x) \mathbf{I}((r-1)/r < x < 1)$ , where  $g(x)$  is a nonnegative function such that  $\int_{(r-1)/r}^1 g(t) dt = 1/2$ , we have  $k = 1$ ,  $f'(0^+) = \frac{(\pi r)^2}{4(r-1)^2}$ , and  $f'((\frac{r-1}{r})^+) = \frac{(\pi r)^2}{4(r-1)^2}$  in Theorem 4.4. Then  $p(F, r, (r-1)/r) = r^2/(r^2 - 1)$ .

(e) For the beta distribution with parameters  $\nu_{1,n}, \nu_{2,n}$ , denoted  $\beta(\nu_{1,n}, \nu_{2,n})$ , where  $\nu_{1,n}, \nu_{2,n} \geq 1$ , the pdf is given by

$$f(x, \nu_{1,n}, \nu_{2,n}) = \frac{x^{\nu_{1,n}-1} (1-x)^{\nu_{2,n}-1}}{\beta(\nu_{1,n}, \nu_{2,n})} \mathbf{I}(0 < x < 1) \text{ where } \beta(\nu_{1,n}, \nu_{2,n}) = \frac{\Gamma(\nu_{1,n}) \Gamma(\nu_{2,n})}{\Gamma(\nu_{1,n} + \nu_{2,n})}.$$

Then in Theorem 4.4 we have  $k = 0$ ,  $f(0^+) = 0$ , and  $f((\frac{r-1}{r})^+) = \frac{(r-1)^{\nu_{1,n}-1}}{r^{\nu_{1,n}+\nu_{2,n}-1}}$ . So  $p(\beta(\nu_{1,n}, \nu_{2,n}), r, (r-1)/r) = 0$ . As for Theorem 4.5, we have  $\ell = 0$ ,  $f(1^-) = 0$ , and  $f((\frac{1}{r})^-) = \frac{(r-1)^{\nu_{2,n}-1}}{r^{\nu_{1,n}+\nu_{2,n}-1}}$ . Then  $p(\beta(\nu_{1,n}, \nu_{2,n}), r, 1/r) = 0$ .

(f) Consider  $F$  with pdf  $f(x) = \left(\pi \sqrt{x(1-x)}\right)^{-1} \mathbf{I}(0 < x < 1)$ . Notice that  $f(x)$  is unbounded at  $x \in \{0, 1\}$ . Using Remark 4.6, it follows that  $p(F, r, (r-1)/r) = p(F, r, 1/r) = 1$ .  $\square$

*Remark 4.9.* In Theorem 4.4, if we have  $f^{(k)}(0^+) = f^{(k)}((\frac{r-1}{r})^+)$ , then  $\lim_{n \rightarrow \infty} p_n(F, r, (r-1)/r) = \frac{1}{1+r^{-(k+1)}}$ . In particular, if  $k = 0$ , then  $\lim_{n \rightarrow \infty} p_n(F, r, (r-1)/r) = r/(r+1)$ . Hence  $\gamma_{n,2}(F, r, (r-1)/r)$  and  $\gamma_{n,2}(\mathcal{U}, r, (r-1)/r)$  have the same asymptotic distribution.

In Theorem 4.5, if we have  $f^{(\ell)}(1^-) = f^{(\ell)}((\frac{1}{r})^-)$ , then  $\lim_{n \rightarrow \infty} p_n(F, r, 1/r) = \frac{1}{1+r^{-(\ell+1)}}$ . In particular, if  $\ell = 0$ , then  $\lim_{n \rightarrow \infty} p_n(F, r, 1/r) = r/(r+1)$ . Hence  $\gamma_{n,2}(F, r, 1/r)$  and  $\gamma_{n,2}(\mathcal{U}, r, 1/r)$  have the same asymptotic distribution.  $\square$

The asymptotic distribution of  $\gamma_{n,2}(F, r, c)$  for  $r = 2$  and  $c = 1/2$  is as follows (see Ceyhan (2008) for its derivation).

**Theorem 4.10.** *Let  $\mathcal{F}(y_1, y_2) := \left\{ F : (y_1, y_1 + \varepsilon) \cup (y_2 - \varepsilon, y_2) \cup ((y_1 + y_2)/2 - \varepsilon, (y_1 + y_2)/2 + \varepsilon) \subseteq \mathcal{S}(F) \subseteq (y_1, y_2) \text{ for some } \varepsilon \in (0, (y_1 + y_2)/2) \right\}$ . Let  $\mathcal{Y}_2 = \{y_1, y_2\} \subset \mathbb{R}$  with  $-\infty < y_1 < y_2 < \infty$ ,  $\mathcal{X}_n = \{X_1, \dots, X_n\}$  with  $X_i \stackrel{iid}{\sim} F \in \mathcal{F}(y_1, y_2)$ , and  $D_{n,2}$  be the random  $\mathcal{D}_{n,2}$ -digraph based on  $\mathcal{X}_n$  and  $\mathcal{Y}_2$ .*

(i) *Then for  $n > 1$ , we have  $\gamma_{n,2}(F, 2, 1/2) \sim 1 + \text{BER}(p_n(F, 2, 1/2))$ . Note also that  $\gamma_{1,2}(F, 2, 1/2) = 1$ .*

(ii) *Furthermore, suppose  $k \geq 0$  is the smallest integer for which  $F(\cdot)$  has continuous right derivatives up to order  $(k+1)$  at  $y_1, (y_1+y_2)/2$ ,  $f^{(k)}(y_1^+) + 2^{-(k+1)} f^{(k)}\left(\left(\frac{y_1+y_2}{2}\right)^+\right) \neq 0$  and  $f^{(i)}(y_1^+) = f^{(i)}\left(\left(\frac{y_1+y_2}{2}\right)^+\right) = 0$  for all  $i = 0, 1, \dots, k-1$ ; and  $\ell \geq 0$  is the smallest integer for which  $F(\cdot)$  has continuous left derivatives up to order  $(\ell+1)$  at  $y_2, (y_1+y_2)/2$ ,  $f^{(\ell)}(y_2^-) + 2^{-(\ell+1)} f^{(\ell)}\left(\left(\frac{y_1+y_2}{2}\right)^-\right) \neq 0$  and  $f^{(i)}(y_2^-) = f^{(i)}\left(\left(\frac{y_1+y_2}{2}\right)^-\right) = 0$  for all  $i = 0, 1, \dots, \ell-1$ . Additionally, for bounded  $f^{(k)}(\cdot)$  and  $f^{(\ell)}(\cdot)$ , we have the following limit*

$$p(F, 2, 1/2) = \lim_{n \rightarrow \infty} p_n(F, 2, 1/2) = \frac{f^{(k)}(y_1^+) f^{(\ell)}(y_2^-)}{\left[ f^{(k)}(y_1^+) + 2^{-(k+1)} f^{(k)}\left(\left(\frac{y_1+y_2}{2}\right)^+\right) \right] \left[ f^{(\ell)}(y_2^-) + 2^{-(\ell+1)} f^{(\ell)}\left(\left(\frac{y_1+y_2}{2}\right)^-\right) \right]}.$$

Notice the interesting behavior of  $p(F, r, c)$  around  $(r, c) = (2, 1/2)$ . There is a jump (hence discontinuity) in  $p(F, r, (r-1)/r)$  and in  $p(F, r, 1/r)$  at  $r = 2$ .

## 5 The Distribution of the Domination Number of $\mathcal{D}_{n,m}(r, c)$ -digraphs

We now consider the more challenging case of  $m > 2$ . For  $\omega_1 < \omega_2$  in  $\mathbb{R}$ , define the family of distributions

$$\mathcal{H}(\mathbb{R}) := \{F_{X,Y} : (X_i, Y_i) \sim F_{X,Y} \text{ with support } \mathcal{S}(F_{X,Y}) = (\omega_1, \omega_2)^2 \subsetneq \mathbb{R}^2, X_i \sim F_X \text{ and } Y_i \stackrel{iid}{\sim} F_Y\}.$$

We provide the exact distribution of  $\gamma_{n,m}(F, r, c)$  for  $\mathcal{H}(\mathbb{R})$ -random digraphs in the following theorem. Let  $[m] := \{0, 1, 2, \dots, m-1\}$  and  $\Theta_{a,b}^S := \{(u_1, u_2, \dots, u_b) : \sum_{i=1}^b u_i = a : u_i \in S, \forall i\}$ . If  $Y_i$  have a continuous distribution, then the order statistics of  $\mathcal{Y}_m$  are distinct a.s. Given  $Y_{(i)} = y_{(i)}$  for  $i = 1, 2, \dots, m$ , let  $\vec{n}$  be the vector of numbers  $n_i$ ,  $f_{\vec{Y}}(\vec{y})$  be the joint distribution of the order statistics of  $\mathcal{Y}_m$ , i.e.,  $f_{\vec{Y}}(\vec{y}) = \frac{1}{m!} \prod_{i=1}^m f(y_i) \mathbf{I}(\omega_1 < y_1 < y_2 < \dots < y_m < \omega_2)$ , and  $f_{i,j}(y_i, y_j)$  be the joint distribution of  $Y_{(i)}, Y_{(j)}$ . Then we have the following theorem.

**Theorem 5.1.** *Let  $D$  be an  $\mathcal{H}(\mathbb{R})$ -random  $\mathcal{D}_{n,m}(r, c)$ -digraph with  $n > 1$ ,  $m > 1$ ,  $r \in [1, \infty)$  and  $c \in (0, 1)$ . Then the probability mass function of the domination number of  $D$  is given by*

$$P(\gamma_{n,m}(F, r, (r-1)/r) = q) = \int_{\mathcal{S}} \sum_{\vec{n} \in \Theta_{n, (m+1)}^{[n+1]}} \sum_{\vec{q} \in \Theta_{q, (m+1)}^{[3]}} P(\vec{N} = \vec{n}) \zeta(q_1, n_1) \zeta(q_{m+1}, n_{m+1}) \prod_{j=2}^m \eta(q_i, n_i) f_{\vec{Y}}(\vec{y}) dy_1 \dots dy_m$$

where  $P(\vec{N} = \vec{n})$  is the joint probability of  $n_i$  points falling into intervals  $\mathcal{I}_i$  for  $i = 0, 1, 2, \dots, m$ ,  $q_i \in \{0, 1, 2\}$ ,  $q = \sum_{i=0}^m q_i$  and

$$\begin{aligned} \zeta(q_i, n_i) &= \max(\mathbf{I}(n_i = q_i = 0), \mathbf{I}(n_i \geq q_i = 1)) \text{ for } i = 1, (m+1), \text{ and} \\ \eta(q_i, n_i) &= \max(\mathbf{I}(n_i = q_i = 0), \mathbf{I}(n_i \geq q_i = 1)) \cdot p_{n_i}(F_i, r, (r-1)/r)^{\mathbf{I}(q_i=2)} (1 - p_{n_i}(F_i, r, (r-1)/r))^{\mathbf{I}(q_i=1)} \\ &\text{for } i = 1, 2, 3, \dots, (m-1), \text{ and the region of integration is given by} \end{aligned}$$

$$\mathcal{S} := \{(y_1, y_2, \dots, y_m) \in (\omega_1, \omega_2)^2 : \omega_1 < y_1 < y_2 < \dots < y_m < \omega_2\}.$$

The special cases of  $n = 1$ ,  $m = 1$ ,  $r \in \{1, \infty\}$  and  $c \in \{0, 1\}$  are as in Theorem 2.4.

Proof is as in Theorem 6.1 of Ceyhan (2008). A similar construction is available for  $c = 1/r$ .

This exact distribution for finite  $n$  and  $m$  has a simpler form when  $\mathcal{X}$  and  $\mathcal{Y}$  points are both uniform in a bounded interval in  $\mathbb{R}$ . Define  $\mathcal{U}(\mathbb{R})$  as follows

$$\mathcal{U}(\mathbb{R}) := \{F_{X,Y} : X \text{ and } Y \text{ are independent } X_i \stackrel{iid}{\sim} \mathcal{U}(\omega_1, \omega_2) \text{ and } Y_i \stackrel{iid}{\sim} \mathcal{U}(\omega_1, \omega_2), \text{ with } -\infty < \omega_1 < \omega_2 < \infty\}.$$

Clearly,  $\mathcal{U}(\mathbb{R}) \subsetneq \mathcal{H}(\mathbb{R})$ . Then we have the following corollary to Theorem 5.1.

**Corollary 5.2.** *Let  $D$  be a  $\mathcal{U}(\mathbb{R})$ -random  $\mathcal{D}_{n,m}(r, c)$ -digraph with  $n > 1$ ,  $m > 1$ ,  $r \in [1, \infty)$  and  $c \in (0, 1)$ . Then the probability mass function of the domination number of  $D$  is given by*

$$P(\gamma_{n,m}(r, (r-1)/r) = q) = \frac{n!m!}{(n+m)!} \sum_{\vec{n} \in \Theta_{n,(m+1)}^{[n+1]}} \sum_{\vec{q} \in \Theta_{q,(m+1)}^{[3]}} \zeta(q_1, n_1) \zeta(q_{m+1}, n_{m+1}) \prod_{j=2}^m \eta(q_j, n_j).$$

The special cases of  $n = 1$ ,  $m = 1$ ,  $r \in \{1, \infty\}$  and  $c \in \{0, 1\}$  are as in Theorem 2.4.

Proof is as in Theorem 2 of Priebe et al. (2001). A similar construction is available for  $c = 1/r$ . For  $n, m < \infty$ , the expected value of domination number is

$$\mathbf{E}[\gamma_{n,m}(F, r, c)] = P(X_{(1)} < Y_{(1)}) + P(X_{(n)} > Y_{(m)}) + \sum_{i=1}^{m-1} \sum_{k=1}^n P(N_i = k) \mathbf{E}[\gamma_{n_i,2}(F_i, r, c)] \quad (15)$$

where

$$P(N_i = k) = \int_{\omega_1}^{\omega_2} \int_{y_{(i)}}^{\omega_2} f_{i-1,i}(y_{(i)}, y_{(i+1)}) \left[ F_X(y_{(i+1)}) - F_X(y_{(i)}) \right]^k \left[ 1 - (F_X(y_{(i+1)}) - F_X(y_{(i)})) \right]^{n-k} dy_{(i+1)} dy_{(i)}$$

and  $\mathbf{E}[\gamma_{n_i,2}(F_i, r, c)] = 1 + p_n(F_i, r, c)$ . Then as in Corollary 6.2 of Ceyhan (2008), we have

**Corollary 5.3.** *For  $F_{X,Y} \in \mathcal{H}(\mathbb{R})$  with support  $\mathcal{S}(F_X) \cap \mathcal{S}(F_Y)$  of positive measure with  $r \in [1, \infty)$  and  $c \in (0, 1)$ , we have  $\lim_{n \rightarrow \infty} \mathbf{E}[\gamma_{n,n}(F, r, c)] = \infty$ .*

**Theorem 5.4. Main Result 5:** *Let  $D_{n,m}(r, c)$  be an  $\mathcal{H}(\mathbb{R})$ -random  $\mathcal{D}_{n,m}(r, c)$ -digraph. Then*

(i) *for fixed  $n < \infty$ ,  $\lim_{m \rightarrow \infty} \gamma_{n,m}(F, r, c) = n$  a.s. for all  $r \geq 1$  and  $c \in [0, 1]$ .*

*For fixed  $m < \infty$ , and*

(ii) *for  $r = 1$  and  $c \in (0, 1)$ ,  $\lim_{n \rightarrow \infty} P(\gamma_{n,m}(F, 1, c) = 2m) = 1$  and  $\lim_{n \rightarrow \infty} P(\gamma_{n,m}(F, 1, 0) = m + 1) = \lim_{n \rightarrow \infty} P(\gamma_{n,m}(F, 1, 1) = m + 1) = 1$*

(iii) *for  $r > 2$  and  $c \in (0, 1)$ ,  $\lim_{n \rightarrow \infty} P(\gamma_{n,m}(F, r, c) = m + 1) = 1$ ,*

(iv) *for  $r = 2$ , if  $c \neq 1/2$ , then  $\lim_{n \rightarrow \infty} P(\gamma_{n,m}(F, 2, c) = m + 1) = 1$ ;*

*if  $c = 1/2$ , then  $\lim_{n \rightarrow \infty} \gamma_{n,m}(F, 2, 1/2) \stackrel{d}{=} m + 1 + \text{BIN}(m, p(F_i, 2, 1/2))$ ,*

(v) *for  $r \in (1, 2)$ , if  $r \neq \tau = \max(c, 1 - c)$ , then  $\lim_{n \rightarrow \infty} \gamma_{n,m}(F, r, c)$  is degenerate; otherwise, it is non-degenerate. That is, for  $r \in [1, 2)$ , as  $n \rightarrow \infty$ ,*

$$\gamma_{n,m}(F, r, c) \sim \begin{cases} m + 1 + \text{BIN}(m, p(F_i, r, c)), & \text{for } r = 1/\tau, \\ m + 1, & \text{for } r > 1/\tau, \\ 2m + 1, & \text{for } r < 1/\tau. \end{cases} \quad (16)$$

**Proof:**

Part (i) is trivial. Part (ii) follows from Theorems 2.4 and 2.5, since as  $n_i \rightarrow \infty$ , we have  $\mathcal{X}_{[i]} \neq \emptyset$  a.s. for all  $i$ .

Part (iii) follows from Theorem 3.6, since for  $c \in (0, 1)$ , it follows that  $r > 1/\tau$  implies  $r > 2$  and as  $n_i \rightarrow \infty$ , we have  $\gamma_{n_i, 2}(F_i, r, c) \rightarrow 1$  in probability for all  $i$ .

In part (iv), for  $r = 2$  and  $c \neq 1/2$ , based on Theorem 3.2, as  $n_i \rightarrow \infty$ , we have  $\gamma_{n_i, 2}(F_i, r, c) \rightarrow 1$  in probability for all  $i$ . The result for  $r = 2$  and  $c = 1/2$  is proved in Ceyhan (2008).

Part (v) follows from Theorem 3.6. ■

**Remark 5.5. Extension to Higher Dimensions:**

Let  $\mathcal{Y}_m = \{y_1, y_2, \dots, y_m\}$  be  $m$  points in general position in  $\mathbb{R}^d$  and  $T_i$  be the  $i^{th}$  Delaunay cell in the Delaunay tessellation (assumed to exist) based on  $\mathcal{Y}_m$  for  $i = 1, 2, \dots, J_m$ . Let  $\mathcal{X}_n$  be a random sample from a distribution  $F$  in  $\mathbb{R}^d$  with support  $\mathcal{S}(F) \subseteq \mathcal{C}_H(\mathcal{Y}_m)$  where  $\mathcal{C}_H(\mathcal{Y}_m)$  stands for the convex hull of  $\mathcal{Y}_m$ . In  $\mathbb{R}$  a Delaunay tessellation is an intervalization (i.e., partitioning of  $\mathbb{R}$  by intervals), provided that no two points in  $\mathcal{Y}_m$  are concurrent.

We define the proportional-edge proximity region in  $\mathbb{R}^2$ . The extension to  $\mathbb{R}^d$  with  $d > 2$  is straightforward (see Ceyhan and Priebe (2007) for an explicit extension). Let  $T(\mathcal{Y}_3)$  be the triangle (including the interior) with vertices  $\mathcal{Y}_3 = \{y_1, y_2, y_3\}$ ,  $e_i$  be the edge opposite vertex  $y_i$ , and  $M_i$  be the midpoint of edge  $e_i$  for  $i = 1, 2, 3$ . We first construct the vertex regions based on a point  $M \in \mathbb{R}^2 \setminus \mathcal{Y}_3$  called *M-vertex regions*, by the lines joining  $M$  to a point on each of the edges of  $T(\mathcal{Y}_3)$ . See Ceyhan (2005) for a more general definition of vertex regions. Preferably,  $M$  is selected to be in the interior of the triangle  $T(\mathcal{Y}_3)$ . For such an  $M$ , the corresponding vertex regions can be defined using a line segment joining  $M$  to  $e_i$ . With center of mass  $M_{CM}$ , the lines joining  $M_{CM}$  and  $\mathcal{Y}_3$  are the *median lines* which cross edges at midpoints  $M_i$  for  $i = 1, 2, 3$ . The vertex regions in Figure 3 are center of mass vertex regions. For  $r \in [1, \infty]$ , define  $N(\cdot, r, M)$  to be the (*parametrized*) *proportional-edge proximity map* with *M-vertex regions* as follows (see also Figure 3 with  $M = M_{CM}$  and  $r = 2$ ). Let  $R_M(v)$  be the vertex region associated with vertex  $v$  and  $M$ . For  $x \in T(\mathcal{Y}_3) \setminus \mathcal{Y}_3$ , let  $v(x) \in \mathcal{Y}_3$  be the vertex whose region contains  $x$ ; i.e.,  $x \in R_M(v(x))$ . If  $x$  falls on the boundary of two *M-vertex regions*, we assign  $v(x)$  arbitrarily. Let  $e(x)$  be the edge of  $T(\mathcal{Y}_3)$  opposite of  $v(x)$ ,  $\ell(v(x), x)$  be the line parallel to  $e(x)$  and passes through  $x$ , and  $d(v(x), \ell(v(x), x))$  be the Euclidean distance from  $v(x)$  to  $\ell(v(x), x)$ . For  $r \in [1, \infty]$ , let  $\ell_r(v(x), x)$  be the line parallel to  $e(x)$  such that

$$d(v(x), \ell_r(v(x), x)) = r d(v(x), \ell(v(x), x)) \text{ and } d(\ell(v(x), x), \ell_r(v(x), x)) < d(v(x), \ell_r(v(x), x)).$$

Let  $T_r(x)$  be the triangle similar to and with the same orientation as  $T(\mathcal{Y}_3)$  having  $v(x)$  as a vertex and  $\ell_r(v(x), x)$  as the opposite edge. Then the *proportional-edge proximity region*  $\mathcal{N}(x, r, M)$  is defined to be  $T_r(x) \cap T(\mathcal{Y}_3)$ . Notice that  $\ell(v(x), x)$  divides the two edges of  $T_r(x)$  (other than the one lies on  $\ell_r(v(x), x)$ ) proportionally with the factor  $r$ . Hence the name *proportional-edge proximity region*.

Notice that in  $\mathbb{R}$ ,  $M$  is the center parametrized by  $c$ , e.g., the center of mass  $M_{CM}$  corresponds to  $c = 1/2$ , but for other  $M \in T(\mathcal{Y}_3)$ , there is no direct counterpart in  $\mathbb{R}$ . The vertex regions in  $\mathbb{R}$  with  $\mathcal{Y}_2 = \{y_1, y_2\}$  are  $(y_1, M_c)$  and  $(M_c, y_2)$ . Observe that  $\mathcal{N}(x, r, c)$  in  $\mathbb{R}$  is an open interval, while in  $\mathbb{R}^d$ , the region  $\mathcal{N}(x, r, M)$  is a closed region. However, the interiors of  $\mathcal{N}(x, r, M)$  satisfy the class cover problem of Cannon and Cowen (2000). □

## 6 Discussion

In this article, we present the distribution of the domination number of a random digraph family called proportional-edge proximity catch digraph (PCD) which is based on two classes of points. Points from one of the classes constitute the vertices of the PCDs, while the points from the other class are used in the binary relation that assigns the arcs of the PCDs.

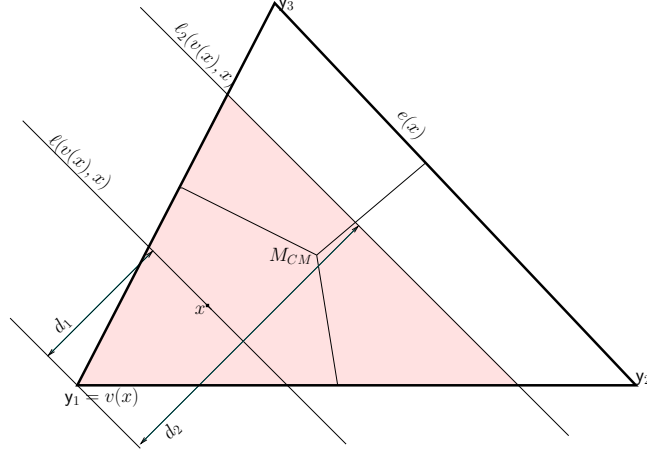


Figure 3: Construction of proportional-edge proximity region,  $\mathcal{N}(x, 2, M_{CM})$  (shaded region) for an  $x$  in the CM-vertex region for  $y_1$ ,  $R_{CM}(y_1)$  where  $d_1 = d(v(x), \ell(v(x), x))$  and  $d_2 = d(v(x), \ell_2(v(x), x)) = 2d(v(x), \ell(v(x), x))$ .

We introduce the proximity map which is the one-dimensional version of  $N(\cdot, r, c)$  of Ceyhan and Priebe (2005) and Ceyhan and Priebe (2007). This proximity map can also be viewed as an extension of the proximity map of Priebe et al. (2001) and Ceyhan (2008). The PCD we consider is based on a parametrized proximity map in which there is an expansion parameter  $r$  and a centrality parameter  $c$ . We provide the exact and asymptotic distributions of the domination number for proportional-edge PCDs for uniform data and compute the asymptotic distribution for non-uniform data for the entire range of  $(r, c)$ . The results in this article can also be viewed as generalizations of the main results of Priebe et al. (2001) and Ceyhan (2008) in several directions. Priebe et al. (2001) provided the exact distribution of the domination number of class cover catch digraphs (CCCDs) based on  $\mathcal{X}_n$  and  $\mathcal{Y}_m$  both of which were sets of iid random variables from uniform distribution on  $(\omega_1, \omega_2) \subset \mathbb{R}$  with  $-\infty < \omega_1 < \omega_2 < \infty$  and the proximity map  $N(x) := B(x, r(x))$  where  $r(x) := \min_{y \in \mathcal{Y}_m} d(x, y)$ . Ceyhan (2008) investigates the distribution of the domination number of CCCDs for non-uniform data and provides the asymptotic distribution for a large family of (non-uniform) distributions. Furthermore, this article will form the foundation of the generalizations and calculations for uniform and non-uniform cases in multiple dimensions. As in Ceyhan and Priebe (2005), we can use the domination number in testing one-dimensional spatial point patterns and our results will help make the power comparisons possible for a large family of distributions.

We demonstrate an interesting behavior of the domination number of proportional-edge PCD for one-dimensional data. For uniform data or data from a distribution which satisfies some regularity conditions (see Section 4.1) and fixed  $1 < n < \infty$ , the distribution of the domination number is a translated form of (extended) binomial distribution  $\text{BIN}(m, p_n(F, r, c))$  where  $m$  is the number of (inner) intervals and  $p_n(F, r, c)$  is the probability that the domination number of the proportional-edge PCD is 2. Here  $p_n(F, r, c)$  is allowed to take 0 or 1 also. For finite  $n > 1$ , the parameter,  $p_n(\mathcal{U}, r, c)$ , of distribution of the domination number of the proportional-edge PCD based on uniform data is continuous in  $r$  and  $c$  for all  $r \geq 1$  and  $c \in (0, 1)$  and has jumps (hence discontinuities) at  $c = 0, 1$ . For fixed  $(r, c) \in [1, \infty) \times (0, 1)$ , the parameter,  $p(\mathcal{U}, r, c)$ , of the asymptotic distribution exhibits some discontinuities. For  $c \in (0, 1)$  the asymptotic distribution is nondegenerate at  $r = 1/\max(c, 1 - c)$ . The asymptotic distribution of the domination number is degenerate for the expansion parameter  $r > 2$ . For  $r \in (1, 2]$ , there exists threshold values for the centrality parameter  $c$  for which the asymptotic distribution is non-degenerate. In particular, for  $c \in \{(r - 1)/r, 1/r\}$  with  $r \in (1, 2]$  the asymptotic distribution of the domination number is a translated form of  $\text{BIN}(m, p(\mathcal{U}, r, c))$  where  $p(F, r, c)$  is continuous in  $r$ . Additionally, by symmetry, we have  $p(\mathcal{U}, r, (r - 1)/r) = p(\mathcal{U}, r, 1/r)$  for  $r \in (1, 2]$ . For  $r > 1/\max(c, 1 - c)$  the domination number converges in probability to 1, and for  $r < 1/\max(c, 1 - c)$  the domination number converges in probability to 2. On the other hand, at  $(r, c) = (2, 1/2)$ , the asymptotic distribution is again a translated form of  $\text{BIN}(m, p(\mathcal{U}, 2, 1/2))$  but there is yet another jump at  $r = 2$ , as  $p(\mathcal{U}, 2, 1/2) = 4/9$  while  $\lim_{r \rightarrow 2} p(\mathcal{U}, r, (r - 1)/r) = \lim_{r \rightarrow 2} p(\mathcal{U}, r, 1/r) = 2/3$ . This second jump might be due

to the symmetry for the domination number at  $c = 1/2$ .

## Acknowledgments

Supported by TUBITAK Kariyer Project Grant 107T647.

## References

- Cannon, A. and Cowen, L. (2000). Approximation algorithms for the class cover problem. In *Proceedings of the 6th International Symposium on Artificial Intelligence and Mathematics*.
- Ceyhan, E. (2005). *An Investigation of Proximity Catch Digraphs in Delaunay Tessellations, also available as technical monograph titled "Proximity Catch Digraphs: Auxiliary Tools, Properties, and Applications" by VDM Verlag, ISBN: 978-3-639-19063-2*. PhD thesis, The Johns Hopkins University, Baltimore, MD, 21218.
- Ceyhan, E. (2008). The distribution of the domination number of class cover catch digraphs for non-uniform one-dimensional data. *Discrete Mathematics*, 308:5376–5393.
- Ceyhan, E. (2010). Spatial clustering tests based on domination number of a new random digraph family. *To appear in Communications in Statistics - Theory and Methods*
- Ceyhan, E. and Priebe, C. (2003). Central similarity proximity maps in Delaunay tessellations. In *Proceedings of the Joint Statistical Meeting, Statistical Computing Section, American Statistical Association*.
- Ceyhan, E. and Priebe, C. E. (2005). The use of domination number of a random proximity catch digraph for testing spatial patterns of segregation and association. *Statistics & Probability Letters*, 73:37–50.
- Ceyhan, E. and Priebe, C. E. (2007). On the distribution of the domination number of a new family of parametrized random digraphs. *Model Assisted Statistics and Applications*, 1(4):231–255.
- Ceyhan, E., Priebe, C. E., and Marchette, D. J. (2007). A new family of random graphs for testing spatial segregation. *Canadian Journal of Statistics*, 35(1):27–50.
- Ceyhan, E., Priebe, C. E., and Wierman, J. C. (2006). Relative density of the random  $r$ -factor proximity catch digraphs for testing spatial patterns of segregation and association. *Computational Statistics & Data Analysis*, 50(8):1925–1964.
- Chartrand, G. and Lesniak, L. (1996). *Graphs & Digraphs*. Chapman & Hall/CRC Press LLC, Florida.
- DeVinney, J. and Priebe, C. E. (2006). A new family of proximity graphs: Class cover catch digraphs. *Discrete Applied Mathematics*, 154(14):1975–1982.
- DeVinney, J., Priebe, C. E., Marchette, D. J., and Socolinsky, D. (2002). Random walks and catch digraphs in classification. <http://www.galaxy.gmu.edu/interface/I02/I2002Proceedings/DeVinneyJason/DeVinneyJason.paper.pdf>. Proceedings of the 34<sup>th</sup> Symposium on the Interface: Computing Science and Statistics, Vol. 34.
- DeVinney, J. and Wierman, J. C. (2003). A SLLN for a one-dimensional class cover problem. *Statistics & Probability Letters*, 59(4):425–435.
- Janson, S., Luczak, T., and Ruciński, A. (2000). *Random Graphs*. Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons, Inc., New York.
- Jaromczyk, J. W. and Toussaint, G. T. (1992). Relative neighborhood graphs and their relatives. *Proceedings of IEEE*, 80:1502–1517.

- Marchette, D. J. and Priebe, C. E. (2003). Characterizing the scale dimension of a high dimensional classification problem. *Pattern Recognition*, 36(1):45–60.
- Paterson, M. S. and Yao, F. F. (1992). On nearest neighbor graphs. In *Proceedings of 19<sup>th</sup> Int. Coll. Automata, Languages and Programming, Springer LNCS*, volume 623, pages 416–426.
- Priebe, C. E., DeVinney, J. G., and Marchette, D. J. (2001). On the distribution of the domination number of random class cover catch digraphs. *Statistics & Probability Letters*, 55:239–246.
- Priebe, C. E., Marchette, D. J., DeVinney, J., and Socolinsky, D. (2003a). Classification using class cover catch digraphs. *Journal of Classification*, 20(1):3–23.
- Priebe, C. E., Solka, J. L., Marchette, D. J., and Clark, B. T. (2003b). Class cover catch digraphs for latent class discovery in gene expression monitoring by DNA microarrays. *Computational Statistics & Data Analysis on Visualization*, 43-4:621–632.
- Prisner, E. (1994). Algorithms for interval catch digraphs. *Discrete Applied Mathematics*, 51:147–157.
- Sen, M., Das, S., Roy, A., and West, D. (1989). Interval digraphs: An analogue of interval graphs. *Journal of Graph Theory*, 13:189–202.
- Toussaint, G. T. (1980). The relative neighborhood graph of a finite planar set. *Pattern Recognition*, 12(4):261–268.
- Tuza, Z. (1994). Inequalities for minimal covering sets in sets in set systems of given rank. *Discrete Applied Mathematics*, 51:187–195.
- West, D. B. (2001). *Introduction to Graph Theory, 2<sup>nd</sup> Edition*. Prentice Hall, NJ.
- Wierman, J. C. and Xiang, P. (2008). A general SLLN for the one-dimensional class cover problem. *Statistics & Probability Letters*, 78(9):1110–1118.
- Xiang, P. and Wierman, J. C. (2009). A CLT for a one-dimensional class cover problem. *Statistics & Probability Letters*, 79(2):223–233.

## APPENDIX

In the proofs of Theorems 3.2, 3.3, 3.4, based on Proposition 3.1, we can assume  $(y_1, y_2) = (0, 1)$ .

### Proof of Theorem 3.2

Given  $X_{(1)} = x_1$  and  $X_{(n)} = x_n$ , let  $a = x_n/2$  and  $b = (1 + x_1)/2$  and due to symmetry, we only consider  $c \in (0, 1/2)$ . For  $r = 2$ , the  $\Gamma_1$ -region is  $\Gamma_1(\mathcal{X}_n, 2, c) = (a, c] \cup [c, b)$  so we have two cases for  $\Gamma_1$ -region: case (1)  $\Gamma_1(\mathcal{X}_n, 2, c) = (a, b)$  which occurs when  $a < c < b$ , and case (2)  $\Gamma_1(\mathcal{X}_n, 2, c) = [c, b)$  which occurs when  $c < a < b$ . The cases in which  $b < a$  and  $b < c$  are not possible, since  $c < 1/2$ ,  $b > 1/2$ , and  $a < 1/2$ .

**Case (1)**  $\Gamma_1(\mathcal{X}_n, 2, c) = (a, b)$ , i.e.,  $a < c < b$ : For  $c \in [1/3, 1/2]$ , we have

$$\begin{aligned}
 P(\gamma_{n,2}(\mathcal{U}, 2, c) = 2, \Gamma_1(\mathcal{X}_n, 2, c) = (a, b), c \in [1/3, 1/2]) = \\
 \left( \int_0^{1/3} \int_{(1+x_1)/2}^{2c} + \int_{1/3}^c \int_{2x_1}^{2c} \right) n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(b)]^{(n-2)} dx_n dx_1 = \\
 \frac{4}{9} \left( 3c - \frac{1}{2} \right)^n - \frac{8}{9} 4^{-n} - \frac{8}{9} \left( \frac{3c-1}{2} \right)^n. \quad (17)
 \end{aligned}$$

For  $c \in (0, 1/3)$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, 2, c) = 2, \Gamma_1(\mathcal{X}_n, 2, c) = (a, b), c \in (0, 1/3)) = \int_0^{4c-1} \int_{(1+x_1)/2}^{2c} n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(b)]^{(n-2)} dx_n dx_1 = \frac{4}{9} (1-3c)^n + \frac{4}{9} \left(3c - \frac{1}{2}\right)^n - \frac{8}{9} 4^{-n}. \quad (18)$$

**Case (2)**  $\Gamma_1(\mathcal{X}_n, 2, c) = [c, b]$ , i.e.,  $c < a < b$ : For  $c \in [1/3, 1/2)$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, 2, c) = 2, \Gamma_1(\mathcal{X}_n, 2, c) = [c, b], c \in [1/3, 1/2)) = \int_0^{4c-1} \int_{(1+x_1)/2}^{2c} n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(c) - F(b)]^{(n-2)} dx_n dx_1 = \frac{2}{3} \left(\frac{3c-1}{2}\right)^n - \frac{2}{3} \left(\frac{1-c}{2}\right)^n - \frac{2}{3} \left(3c - \frac{1}{2}\right)^n + \frac{2}{3} \left(c + \frac{1}{2}\right)^n. \quad (19)$$

For  $c \in (0, 1/3)$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, 2, c) = 2, \Gamma_1(\mathcal{X}_n, 2, c) = [c, b], c \in (0, 1/3)) = \left( \int_0^{4c-1} \int_{2c}^1 + \int_{4c-1}^c \int_{(1+x_1)/2}^1 \right) n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(c) - F(b)]^{(n-2)} dx_n dx_1 = \frac{2}{3} \left(c + \frac{1}{2}\right)^n - \frac{1}{3} (1-3c)^n - \frac{2}{3} \left(3c - \frac{1}{2}\right)^n - \frac{2}{3} \left(\frac{1-c}{2}\right)^n \quad (20)$$

The probability  $P(\gamma_{n,2}(\mathcal{U}, 2, c) = 2, c \in [1/3, 1/2))$  is the sum of probabilities in (17) and (19), and  $P(\gamma_{n,2}(\mathcal{U}, 2, c) = 2, c \in (0, 1/3))$  is the sum of probabilities in (18) and (20). By symmetry,  $P(\gamma_{n,2}(\mathcal{U}, 2, c) = 2, c \in [1/2, 1)) = P(\gamma_{n,2}(\mathcal{U}, 2, 1-c) = 2, 1-c \in (0, 1/2])$ . The special cases for  $c = \{0, 1\}$  follow by construction. ■

### Proof of Theorem 3.3

Given  $X_{(1)} = x_1$  and  $X_{(n)} = x_n$ , let  $a = x_n/r$  and  $b = (x_1 + r - 1)/r$ . For  $r \geq 1$  and  $c = 1/2$ , the  $\Gamma_1$ -region is  $\Gamma_1(\mathcal{X}_n, r, 1/2) = (a, 1/2] \cup [1/2, b)$ , so we have four cases for  $\Gamma_1$ -region: case (1)  $\Gamma_1(\mathcal{X}_n, r, 1/2) = (a, b)$  which occurs when  $a < 1/2 < b$ , case (2)  $\Gamma_1(\mathcal{X}_n, r, 1/2) = (a, 1/2]$  which occurs when  $a < b < 1/2$  or  $b < a < 1/2$ , case (3)  $\Gamma_1(\mathcal{X}_n, r, 1/2) = [1/2, b)$  which occurs when  $1/2 < a < b$  or  $1/2 < b < a$ , and case (4)  $\Gamma_1(\mathcal{X}_n, r, 1/2) = \emptyset$  which occurs when  $b < 1/2 < a$ . Cases (2) and (3) are symmetric, so they yield the same probabilities.

**Case (1)**  $\Gamma_1(\mathcal{X}_n, r, 1/2) = (a, b)$ , i.e.,  $a < 1/2 < b$ : For  $r \geq 2$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, 1/2) = 2, \Gamma_1(\mathcal{X}_n, r, 1/2) = (a, b), r \geq 2) = \left( \int_0^{1/(r+1)} \int_{(x_1+r-1)/r}^1 + \int_{1/(r+1)}^{1/r} \int_{rx_1}^1 \right) n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(b)]^{(n-2)} dx_n dx_1 = \frac{2r}{(r+1)^2} \left[ \left(\frac{2}{r}\right)^{n-1} - \left(\frac{r-1}{r^2}\right)^{n-1} \frac{1}{r} \right]. \quad (21)$$

For  $1 \leq r < 2$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, 1/2) = 2, \Gamma_1(\mathcal{X}_n, r, 1/2) = (a, b), 1 \leq r < 2) =$$

$$\left( \int_{1-r/2}^{1/(r+1)} \int_{(x_1+r-1)/r}^{r/2} + \int_{1/(r+1)}^{1/2} \int_{rx_1}^{r/2} \right) n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(b)]^{(n-2)} dx_n dx_1 =$$

$$\frac{r^2(r-1)^n}{(r+1)^2} \left[ 1 - \left( \frac{r-1}{2r} \right)^{n-1} \right]. \quad (22)$$

**Case (2)**  $\Gamma_1(\mathcal{X}_n, r, 1/2) = (a, 1/2]$ , i.e.,  $a < b < 1/2$  or  $b < a < 1/2$ : Here,  $r \geq 2$  is not possible, since  $x_1 < 1 - r/2$ . For  $1 \leq r < 2$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, 1/2) = 2, \Gamma_1(\mathcal{X}_n, r, 1/2) = (a, 1/2], 1 \leq r < 2) =$$

$$\int_0^{1-r/2} \int_{1/2}^{r/2} n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(1/2)]^{(n-2)} dx_n dx_1 =$$

$$\frac{r}{(r+1)} \left[ \left( \frac{r}{2} \right)^n - (r-1)^n - \left( \frac{1}{2r} \right)^n + \left( \frac{(r-1)^2}{2r} \right)^n \right]. \quad (23)$$

By symmetry, **Case (3)** yields the same result as **Case (2)**.

**Case (4)**  $\Gamma_1(\mathcal{X}_n, r, 1/2) = \emptyset$ , i.e.,  $b < 1/2 < a$ : Here,  $r \geq 2$  is not possible, since  $x_n - x_1 > r - 1$ . For  $1 \leq r < 2$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, 1/2) = 2, \Gamma_1(\mathcal{X}_n, r, 1/2) = \emptyset, 1 < r < 2) =$$

$$\int_0^{1-r/2} \int_{1/2}^1 n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1)]^{(n-2)} dx_n dx_1 = 1 + (r-1)^n - 2 \left( \frac{r}{2} \right)^n. \quad (24)$$

The probability  $P(\gamma_{n,2}(\mathcal{U}, r, 1/2) = 2, r \geq 2)$  is as in (21), and  $P(\gamma_{n,2}(\mathcal{U}, r, 1/2) = 2, r \in [1, \infty))$  is the sum of probabilities in (22), (24), and twice the probability in (23). ■

### Proof of Theorem 3.4

Given  $X_{(1)} = x_1$  and  $X_{(n)} = x_n$ , let  $a = x_n/r$  and  $b = (x_1 + r - 1)/r$  and assume  $c \in (0, 1/2)$ . There are two cases for  $c$ , namely, **Case I-**  $c \in ((3 - \sqrt{5})/2, 1/2)$  and **Case II-**  $c \in (0, (3 - \sqrt{5})/2]$

**Case I-** For  $r \geq 1$  and  $c \in ((3 - \sqrt{5})/2, 1/2)$ , the  $\Gamma_1$ -region is  $\Gamma_1(\mathcal{X}_n, r, c) = (a, c] \cup [c, b)$ . So we have four cases for  $\Gamma_1$ -region: case (1)  $\Gamma_1(\mathcal{X}_n, r, c) = (a, b)$  which occurs when  $a < c < b$ , case (2)  $\Gamma_1(\mathcal{X}_n, r, c) = (a, c]$  which occurs when  $a < b < c$  or  $b < a < c$ , case (3)  $\Gamma_1(\mathcal{X}_n, r, c) = [c, b)$  which occurs when  $c < a < b$  or  $c < b < a$ , and case (4)  $\Gamma_1(\mathcal{X}_n, r, c) = \emptyset$  which occurs when  $b < c < a$ .

**Case (1)**  $\Gamma_1(\mathcal{X}_n, r, c) = (a, b)$ , i.e.,  $a < c < b$ : In this case, for  $r \geq 1/c$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = (a, b), r \geq 1/c) =$$

$$\left( \int_0^{1/(r+1)} \int_{(x_1+r-1)/r}^1 + \int_{1/(r+1)}^{1/r} \int_{rx_1}^1 \right) n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(b)]^{(n-2)} dx_n dx_1 =$$

$$\frac{2r}{(r+1)^2} \left( \left( \frac{2}{r} \right)^{n-1} - \left( \frac{r-1}{r^2} \right)^{n-1} \right) \quad (25)$$

For  $1/(1-c) \leq r < 1/c$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = (a, b), 1/(1-c) \leq r < 1/c) =$$

$$\left( \int_0^{1/(r+1)} \int_{(x_1+r-1)/r}^{cr} + \int_{1/(r+1)}^c \int_{rx_1}^{cr} \right) n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(b)]^{(n-2)} dx_n dx_1 =$$

$$\frac{r^2}{(r+1)^2} \left[ \left( c(r+1) - \frac{r-1}{r} \right)^n - \left( \frac{r-1}{r} \right)^{n-1} \left( (cr+c-1)^n - \frac{1}{r^n} \right) \right]. \quad (26)$$

For  $(1-c)/c \leq r < 1/(1-c)$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = (a, b), (1-c)/c \leq r < 1/(1-c)) =$$

$$\left( \int_{r(c-1)+1}^{1/(r+1)} \int_{(x_1+r-1)/r}^{cr} + \int_{1/(r+1)}^c \int_{rx_1}^{cr} \right) n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(b)]^{(n-2)} dx_n dx_1 =$$

$$\frac{r^2(r-1)^{n-1}}{(r+1)^2} \left[ (r-1) - \frac{1}{r^{n-1}} [(r-cr-c)^n - (cr+c-1)^n] \right]. \quad (27)$$

For  $1 \leq r < (1-c)/c$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = (a, b), 1 \leq r < (1-c)/c) =$$

$$\int_{r(c-1)+1}^{cr^2-r+1} \int_{(x_1+r-1)/r}^{cr} n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(b)]^{(n-2)} dx_n dx_1 =$$

$$\frac{r^2(r-1)^{n-1}}{(r+1)^2} \left[ r-1 + (1-cr-c)^n + \frac{(r-cr-c)^n}{r^{n-1}} \right]. \quad (28)$$

**Case (2)**  $\Gamma_1(\mathcal{X}_n, r, c) = (a, c]$ , i.e.,  $a < b < c$  or  $b < a < c$ : Here  $r \geq 1/(1-c)$  is not possible, since  $r(c-1)+1 > 0$ . Hence  $r \geq 1/c$  is not possible either. For  $1 \leq r < 1/(1-c)$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = (a, c], 1 \leq r < 1/(1-c)) =$$

$$\int_0^{r(c-1)+1} \int_c^{cr} n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(c)]^{(n-2)} dx_n dx_1 =$$

$$\frac{r}{r+1} \left[ c^n \left( r^n - \frac{1}{r^n} \right) - (r-1)^n \left( 1 - \frac{r-cr-c}{r} \right)^n \right]. \quad (29)$$

**Case (3)**  $\Gamma_1(\mathcal{X}_n, r, c) = [c, b)$ , i.e.,  $c < a < b$  or  $c < b < a$ : Here  $r \geq 1/c$  is not possible, since  $x_n > cr$ . For  $1/(1-c) \leq r < 1/c$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = [c, b), 1/(1-c) \leq r < 1/c) =$$

$$\int_0^c \int_{cr}^1 n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(c) - F(b)]^{(n-2)} dx_n dx_1 =$$

$$\frac{1}{(r+1)r^{n-1}} \left[ (r-1)^n (cr-1+c)^n + (1+cr)^n - (cr^2+cr-r+1)^n - (1-c)^n \right]. \quad (30)$$

For  $1 \leq r < 1/(1-c)$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = [c, b), 1 \leq r < 1/(1-c)) =$$

$$\int_{r(c-1)+1}^c \int_{cr}^1 n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(c) - F(b)]^{(n-2)} dx_n dx_1 =$$

$$\frac{r}{r+1} \left[ (r-1)^n \left( \left( \frac{cr-1+c}{r} \right)^n - 1 \right) + (1-c)^n \left( r^n - \frac{1}{r^n} \right) \right]. \quad (31)$$

**Case (4)**  $\Gamma_1(\mathcal{X}_n, r, c) = \emptyset$ , i.e.,  $b < c < a$ : Here,  $r \geq 1/c$  is not possible, since  $x_n > cr$ ; and  $1/(1-c) \leq r < 1/c$  is not possible, since  $x_1 < r(c-1) + 1$ . For  $1 \leq r < 1/(1-c)$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = \emptyset, 1 \leq r < 1/(1-c)) = \int_0^{r(c-1)+1} \int_{cr}^1 n(n-1)f(x_1)f(x_n)(F(x_n) - F(x_1))^{(n-2)} dx_n dx_1 = 1 + (r-1)^n - r^n[c^n + (1-c)^n]. \quad (32)$$

The probability  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, r \geq 1/c)$  is the same as in (25);  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, 1/(1-c) \leq r < 1/c)$  is the sum of probabilities in (26) and (30);  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, (1-c)/c \leq r < 1/(1-c))$  is the sum of probabilities in (27), (29), (31), and (32);  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, 1 \leq r < (1-c)/c)$  is the sum of probabilities in (28), (29), (31), and (32).

**Case II-** For  $r \geq 1$  and  $c \in (0, (3 - \sqrt{5})/2)$ , we have the same cases for the  $\Gamma_1$ -region is as above.

**Case (1):** For  $r \geq 1/c$ , the probability  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = (a, b), r \geq 1/c)$  as in (25).

For  $(1-c)/c \leq r < 1/c$ , the probability  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = (a, b), (1-c)/c \leq r < 1/c)$  is as in (26).

For  $1/(1-c) \leq r < (1-c)/c$ , we have

$$P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = (a, b), 1/(1-c) \leq r < (1-c)/c) = \int_0^{cr^2-r+1} \int_{(x_1+r-1)/r}^{cr} n(n-1)f(x_1)f(x_n)[F(x_n) - F(x_1) + F(a) - F(b)]^{(n-2)} dx_n dx_1 = \frac{r^2}{(r+1)^2} \left[ (r-1)^{n-1} \left( (1-cr-c)^n - \frac{1}{r^{2n-1}} \right) + \left( \frac{cr^2+cr-r+1}{r} \right)^n \right]. \quad (33)$$

For  $1 \leq r < 1/(1-c)$ , the probability  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = (a, b), 1 \leq r < 1/(1-c))$  is as in (28).

**Cases (2) and (4)** are as before.

**Case (3):** For  $(1-c)/c \leq r < 1/c$ , the probability  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = [c, b], (1-c)/c \leq r < 1/c)$  is as in (30).

For  $1/(1-c) \leq r < (1-c)/c$ , the probability  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = [c, b], 1/(1-c) \leq r < (1-c)/c)$  is as in (30).

For  $1 \leq r < 1/(1-c)$ , the probability  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, \Gamma_1(\mathcal{X}_n, r, c) = [c, b], 1 \leq r < 1/(1-c))$  is as in (31).

The probability  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, r \geq 1/c)$  is the same as in (25);  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, (1-c)/c \leq r < 1/c)$  is the sum of probabilities in (26) and (30);  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, 1/(1-c) \leq r < (1-c)/c)$  is the sum of probabilities in (30) and (33);  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, 1 \leq r < 1/(1-c))$  is the sum of probabilities in (28), (29), (31), and (32).

By symmetry,  $P(\gamma_{n,2}(\mathcal{U}, r, c) = 2, c \in [1/2, 1)) = P(\gamma_{n,2}(\mathcal{U}, r, 1-c) = 2, 1-c \in (0, 1/2])$ . The special case for  $c \in \{0, 1\}$  follows trivially by construction. ■

## Proof of Theorem 3.6

Let  $c \in (0, 1/2)$ . Then  $\tau = 1 - c$ . We first consider **Case I**:  $c \in ((3 - \sqrt{5})/2, 1/2)$ . In Theorem 3.4, for  $r \geq 1/c > 2$ , it follows that  $\lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, c) = \lim_{n \rightarrow \infty} \pi_{1,n}(r, c) = 0$ , since  $2/r < 1$  and  $\frac{r-1}{r^2} < 1$ .

For  $1/(1-c) < r < 1/c$ , we have  $\frac{1+cr}{r} < 1$  (since  $r > 1/(1-c)$ ),  $\frac{1-c}{r} < 1$ ,  $\frac{(cr^2-r+cr+1)}{r} < 1$ ,  $\frac{r-1}{r^2} < 1$  (since  $r-1 < r < r^2$ ), and  $\frac{(r-1)(cr-1+c)}{r} < 1$ . Hence for  $1/(1-c) < r < 1/c$   $\lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, c) = \lim_{n \rightarrow \infty} \pi_{2,n}(r, c) = 0$

For  $(1-c)/c < r < 1/(1-c)$ , we have  $r-1 < 1$  (since  $r < 1/(1-c) < 2$ ),  $\frac{(r-1)(cr-1+c)}{r} < 1$ ,  $\frac{(r-1)(r-cr-c)}{r} < 1$ ,  $c/r < 1$  (since  $c < r$ ),  $(1-c)/r < 1$  (since  $1-c < r$ ),  $cr < 1$  and  $(1-c)r < 1$  (since  $r < 1/(1-c) < 1/c$ ). Hence  $\lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, c) = \lim_{n \rightarrow \infty} \pi_{3,n}(r, c) = 1$ .

For  $1 \leq r < (1-c)/c$ , we have  $r-1 < 1$  (since  $r < 1/(1-c) < 2$ ),  $(r-1)(1-cr-c) < 1$ ,  $\frac{(r-1)(1-cr-c)}{r} < 1$ ,  $\frac{(r-1)(r-cr-c)}{r} < 1$ ,  $c/r < 1$  (since  $c < r$ ),  $(1-c)/r < 1$ ,  $cr < 1$  and  $(1-c)r < 1$ . Hence  $\lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, c) = \lim_{n \rightarrow \infty} \pi_{4,n}(r, c) = 1$ ,

But for  $r = 1/(1-c)$  or  $c = (r-1)/r$ , we have

$$p_n(\mathcal{U}, r, (r-1)/r) = \frac{r}{(r+1)^2} \left[ (r+1) - (r-1)^{n-1} \left( \frac{r^2-r-1}{r^2} \right)^n + (r-1)^n - \frac{r+1}{r^{2n}} - \left( \frac{r-1}{r^2} \right)^{n-1} \right]. \quad (34)$$

Letting  $n \rightarrow \infty$ , we get  $p_n(\mathcal{U}, r, (r-1)/r) \rightarrow r/(r+1)$  for  $r \in (1, 2)$ , since  $\frac{r-1}{r^2} < 1$ ,  $r-1 < 1$ ,  $\frac{1}{r^2} < 1$ , and  $\frac{(r-1)(r^2-r-1)}{r^2} < 1$ .

Next we consider **Case II**:  $c \in (0, (3-\sqrt{5})/2]$ . In Theorem 3.4, for  $r \geq 1/c > 2$ , it follows that  $\lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, c) = \lim_{n \rightarrow \infty} \vartheta_{1,n}(r, c) = 0$ , since  $2/r < 1$  and  $\frac{r-1}{r^2} < 1$ .

For  $(1-c)/c < r < 1/c$ , we have  $\frac{1+cr}{r} < 1$  (since  $r > 1/(1-c)$ ),  $\frac{1-c}{r} < 1$ ,  $\frac{(cr^2-r+cr+1)}{r} < 1$ ,  $\frac{r-1}{r} < 1$  (since  $r-1 < r < r^2$ ), and  $\frac{(r-1)(cr-1+c)}{r} < 1$ . Hence  $\lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, c) = \lim_{n \rightarrow \infty} \vartheta_{2,n}(r, c) = 0$

For  $1/(1-c) < r < (1-c)/c$ , we have  $(r-1)(1-cr-c) < 1$ ,  $\frac{(r-1)(1-cr-c)}{r} < 1$ ,  $\frac{r-1}{r} < 1$ ,  $\frac{r-1}{r^2} < 1$ ,  $(1+cr)/r < 1$ ,  $\frac{cr^2-c+cr+1}{r} < 1$ ,  $c/r < 1$  (since  $c < r$ ),  $(1-c)/r < 1$  (since  $1-c < r$ ),  $cr < 1$  and  $(1-c)r < 1$  (since  $r < 1/(1-c) < 1/c$ ). Hence  $\lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, c) = \lim_{n \rightarrow \infty} \vartheta_{3,n}(r, c) = 0$ .

For  $1 \leq r < 1/(1-c)$ , we have  $r-1 < 1$  (since  $r < 1/(1-c) < 2$ ),  $(r-1)(1-cr-c) < 1$ ,  $\frac{(r-1)(1-cr-c)}{r} < 1$ ,  $\frac{(r-1)(r-cr-c)}{r} < 1$ ,  $c/r < 1$  (since  $c < r$ ),  $(1-c)/r < 1$ ,  $cr < 1$  and  $(1-c)r < 1$ . Hence  $\lim_{n \rightarrow \infty} p_n(\mathcal{U}, r, c) = \lim_{n \rightarrow \infty} \vartheta_{4,n}(r, c) = 1$ ,

But for  $r = 1/(1-c)$  or  $c = (r-1)/r$ , we have

$$p_n(\mathcal{U}, r, (r-1)/r) = \frac{r}{(r+1)^2} \left[ (r+1) + (r+1) \left( \frac{(r-1)(r^2-r-1)}{r^2} \right)^n - (r-1)^n + (-1)^n \left( \frac{r-1}{r} \right)^{n-1} (r^2-r-1)^n - \frac{r-1}{r^{2n}} + \left( \frac{r-1}{r^2} \right)^{n-1} \right]. \quad (35)$$

Letting  $n \rightarrow \infty$ , we get  $p_n(\mathcal{U}, r, (r-1)/r) \rightarrow r/(r+1)$  for  $r \in (1, 2)$ , since  $r-1 < 1$ ,  $\frac{1}{r^2} < 1$ ,  $\frac{(r-1)(r^2-r-1)}{r^2} < 1$ , and  $r^2-r-1 < 1$ .

For  $c \in (1/2, 1)$ , we have  $\tau = c$ . By symmetry, the above results follow with  $c$  being replaced by  $1-c$  and as  $n \rightarrow \infty$ , we get  $p_n(\mathcal{U}, r, 1/r) \rightarrow r/(r+1)$  for  $r \in (1, 2)$ . Hence the desired result follows. ■

## Proof of Theorem 4.4

Case (i) follows trivially from Theorem 2.3. The special cases for  $n = 1$  and  $r = \{1, \infty\}$  follow by construction.

Case (ii): Suppose  $(y_1, y_2) = (0, 1)$  and  $c \in (0, 1/2)$ . Recall that  $\Gamma_1(\mathcal{X}_n, r, c) = (X_{(n)}/r, M_c] \cup [M_c, (X_{(1)} + r - 1)/r) \subset (0, 1)$  and  $\gamma_{n,2}(F, r, c) = 2$  iff  $\mathcal{X}_n \cap \Gamma_1(\mathcal{X}_n, r, c) = \emptyset$ . Then for finite  $n$ ,

$$p_n(F, r, c) = P(\gamma_{n,2}(F, r, c) = 2) = \int_{\mathcal{S}(F) \setminus (\delta_1, \delta_2)} H(x_1, x_n) dx_n dx_1,$$

where  $(\delta_1, \delta_2) = \Gamma_1(\mathcal{X}_n, r, c)$  and  $H(x_1, x_n)$  is as in Equation (12).

Let  $\varepsilon \in (0, (r-1)/r)$  and  $c = (r-1)/r$ . Then  $P(X_{(1)} < \varepsilon, X_{(n)} > 1 - \varepsilon) \rightarrow 1$  as  $n \rightarrow \infty$  with the rate of convergence depending on  $F$ . Moreover, for sufficiently large  $n$ ,  $(X_{(1)} + r - 1)/r > (r-1)/r$  a.s.; in fact,  $(X_{(1)} + r - 1)/r \downarrow (r-1)/r$  as  $n \rightarrow \infty$  (in probability) and  $X_{(n)}/r > \max((r-1)/r, (X_{(1)} + r - 1)/r)$  a.s. since  $r \in (1, 2)$ . Then for sufficiently large  $n$ , we have  $\Gamma_1(\mathcal{X}_n, r, (r-1)/r) = [(r-1)/r, (X_{(1)} + r - 1)/r)$  a.s. and

$$\begin{aligned} p_n(F, r, (r-1)/r) &\approx \int_0^\varepsilon \int_{1-\varepsilon}^1 n(n-1) f(x_1) f(x_n) \left[ F(x_n) - F(x_1) + F((r-1)/r) - F((x_1 + r - 1)/r) \right]^{n-2} dx_n dx_1 \\ &= \int_0^\varepsilon n f(x_1) \left( \left[ 1 - F(x_1) + F((r-1)/r) - F((x_1 + r - 1)/r) \right]^{n-1} - \right. \\ &\quad \left. \left[ 1 - \varepsilon - F(x_1) + F((r-1)/r) - F((x_1 + r - 1)/r) \right]^{n-1} \right) dx_1 \\ &\approx \int_0^\varepsilon n f(x_1) \left[ 1 - F(x_1) + F((r-1)/r) - F((x_1 + r - 1)/r) \right]^{n-1} dx_1. \end{aligned} \quad (36)$$

Let

$$G(x_1) = 1 - F(x_1) + F((r-1)/r) - F((x_1 + r - 1)/r).$$

The integral in Equation (36) is critical at  $x_1 = 0$ , since  $G(0) = 1$ , and for  $x_1 \in (0, 1)$  the integral converges to 0 as  $n \rightarrow \infty$ . Let  $\alpha_i := -\frac{d^{i+1}G(x_1)}{dx_1^{i+1}} \Big|_{(0^+, 0^+)} = f^{(i)}(0^+) + r^{-(i+1)} f^{(i)}\left(\left(\frac{r-1}{r}\right)^+\right)$ . Then by the hypothesis of the theorem, we have  $\alpha_i = 0$  and  $f^{(i)}\left(\left(\frac{r-1}{r}\right)^+\right) = 0$  for all  $i = 0, 1, 2, \dots, (k-1)$ . So the Taylor series expansions of  $f(x_1)$  around  $x_1 = 0^+$  up to order  $k$  and  $G(x_1)$  around  $0^+$  up to order  $(k+1)$  so that  $x_1 \in (0, \varepsilon)$ , are as follows:

$$f(x_1) = \frac{1}{k!} f^{(k)}(0^+) x_1^k + O(x_1^{k+1})$$

and

$$G(x_1) = G(0^+) + \frac{1}{(k+1)!} \left( \frac{d^{k+1}G(0^+)}{dx_1^{k+1}} \right) x_1^{k+1} + O(x_1^{k+2}) = 1 - \frac{\alpha_k}{(k+1)!} x_1^{k+1} + O(x_1^{k+2}).$$

Then substituting these expansions in Equation (36), we obtain

$$p_n(F, r, (r-1)/r) \approx \int_0^\varepsilon n \left[ \frac{1}{k!} f^{(k)}(0^+) x_1^k + O(x_1^{k+1}) \right] \left[ 1 - \frac{\alpha_k}{(k+1)!} x_1^{k+1} + O(x_1^{k+2}) \right]^{n-1} dx_1.$$

Now we let  $x_1 = w n^{-1/(k+1)}$  to obtain

$$\begin{aligned} p_n(F, r, (r-1)/r) &\approx \int_0^{\varepsilon n^{1/(k+1)}} n \left[ \frac{1}{n^{k/(k+1)} k!} f^{(k)}(0^+) w^k + O(n^{-1}) \right] \\ &\quad \left[ 1 - \frac{1}{n} \left( \frac{\alpha_k}{(k+1)!} w^{k+1} + O(n^{-(k+2)/(k+1)}) \right) \right]^{n-1} \left( \frac{1}{n^{1/(k+1)}} \right) dw \end{aligned}$$

letting  $n \rightarrow \infty$ ,

$$\approx \int_0^\infty \frac{1}{k!} f^{(k)}(0^+) w^k \exp \left[ -\frac{\alpha_k}{(k+1)!} w^{k+1} \right] dw = \frac{f^{(k)}(0^+)}{\alpha_k} = \frac{f^{(k)}(0^+)}{f^{(k)}(0^+) + r^{-(k+1)} f^{(k)}\left(\left(\frac{r-1}{r}\right)^+\right)}, \quad (37)$$

as  $n \rightarrow \infty$  at rate  $O(\kappa_1(f) \cdot n^{-(k+2)/(k+1)})$ .

For the general case of  $\mathcal{Y} = \{y_1, y_2\}$ , the transformation  $\phi(x) = (x - y_1)/(y_2 - y_1)$  maps  $(y_1, y_2)$  to  $(0, 1)$  and the transformed random variables  $U = \phi(X_i)$  are distributed with density  $g(u) = (y_2 - y_1)f(y_1 + u(y_2 - y_1))$  on  $(y_1, y_2)$ . Replacing  $f(x)$  by  $g(x)$  in Equation (37), the desired result follows. ■

## Proof of Theorem 4.5

Case (i) follows trivially from Theorem 2.3. The special cases for  $n = 1$  and  $r = \{1, \infty\}$  follow by construction.

Case (ii): Suppose  $(y_1, y_2) = (0, 1)$  and  $c \in (1/2, 1)$ . Let  $\varepsilon \in (0, 1/r)$ . Then  $P(X_{(1)} < \varepsilon, X_{(n)} > 1 - \varepsilon) \rightarrow 1$  as  $n \rightarrow \infty$  with the rate of convergence depending on  $F$ . Moreover, for sufficiently large  $n$ ,  $X_{(n)}/r < 1/r$  a.s.; in fact,  $X_{(n)}/r \uparrow 1/r$  as  $n \rightarrow \infty$  (in probability) and  $(X_{(1)} + r - 1)/r < \min(1/r, X_{(n)}/r)$  a.s. Then for sufficiently large  $n$ ,  $\Gamma_1(\mathcal{X}_n, r, c) = (X_{(n)}/r, 1/r]$  a.s. and

$$\begin{aligned} p_n(F, r, 1/r) &\approx \int_{1-\varepsilon}^1 \int_0^\varepsilon n(n-1)f(x_1)f(x_n) \left[ F(x_n) - F(x_1) + F(x_n/r) - F(1/r) \right]^{n-2} dx_1 dx_n. \\ &= - \int_{1-\varepsilon}^1 n f(x_n) \left( \left[ F(x_n) - F(\varepsilon) + F(x_n/r) - F(1/r) \right]^{n-1} - \left[ F(x_n) + F(x_n/r) - F(1/r) \right]^{n-1} \right) dx_n \\ &\approx \int_{1-\varepsilon}^1 n f(x_n) \left[ F(x_n) + F(x_n/r) - F(1/r) \right]^{n-1} dx_n. \end{aligned} \quad (38)$$

Let

$$G(x_n) = F(x_n) + F(x_n/r) - F(1/r).$$

The integral in Equation (38) is critical at  $x_n = 1$ , since  $G(1) = 1$ , and for  $x_n \in (0, 1)$  the integral converges to 0 as  $n \rightarrow \infty$ . So we make the change of variables  $z_n = 1 - x_n$ , then  $G(x_n)$  becomes

$$G(z_n) = F(1 - z_n) + F((1 - z_n)/r) - F(1/r),$$

and Equation (38) becomes

$$p_n(F, r, 1/r) \approx \int_0^\varepsilon n f(1 - z_n) [G(z_n)]^{n-1} dz_n. \quad (39)$$

The new integral is critical at  $z_n = 0$ . Let  $\beta_i := (-1)^{i+1} \frac{d^{i+1}G(z_n)}{dz_n^{i+1}} \Big|_{0^+} = f^{(i)}(1^-) + r^{-(i+1)} f^{(i)}\left(\left(\frac{1}{r}\right)^-\right)$ . Then by the hypothesis of the theorem, we have  $\beta_i = 0$  and  $f^{(i)}\left(\left(\frac{1}{r}\right)^-\right) = 0$  for all  $i = 0, 1, 2, \dots, (\ell - 1)$ . So the Taylor series expansions of  $f(1 - z_n)$  around  $z_n = 0^+$  up to  $\ell$  and  $G(z_n)$  around  $0^+$  up to order  $(\ell + 1)$  so that  $z_n \in (0, \varepsilon)$ , are as follows:

$$\begin{aligned} f(1 - z_n) &= \frac{(-1)^\ell}{\ell!} f^{(\ell)}(1^-) z_n^\ell + O(z_n^{\ell+1}) \\ G(z_n) &= G(0^+) + \frac{1}{(\ell+1)!} \left( \frac{d^{\ell+1}G(0^+)}{dz_n^{\ell+1}} \right) z_n^{\ell+1} + O(z_n^{\ell+2}) = 1 + \frac{(-1)^{\ell+1} \beta_\ell}{(\ell+1)!} z_n^{\ell+1} + O(z_n^{\ell+2}). \end{aligned}$$

Then substituting these expansions in Equation (39), we get

$$p_n(F, r, 1/r) \approx \int_0^\varepsilon n \left[ \frac{(-1)^\ell}{\ell!} f^{(\ell)}(1^-) z_n^\ell + O(z_n^{\ell+1}) \right] \left[ 1 - \frac{(-1)^\ell \beta_\ell}{(\ell+1)!} z_n^{\ell+1} + O(z_n^{\ell+2}) \right]^{n-1} dz_n.$$

Now we let  $z_n = v n^{-1/(\ell+1)}$ , to obtain

$$\begin{aligned}
p_n(F, r, 1/r) &\approx \int_0^{\varepsilon n^{1/(\ell+1)}} n \left[ \frac{(-1)^\ell}{n^{\ell/(\ell+1)} \ell!} f^{(\ell)}(1^-) v^\ell + O(n^{-1}) \right] \\
&\quad \left[ 1 - \frac{1}{n} \left( \frac{(-1)^\ell \beta_\ell}{(\ell+1)!} v^{\ell+1} + O(n^{-(\ell+2)/(\ell+1)}) \right) \right]^{n-1} \left( \frac{1}{n^{1/(\ell+1)}} \right) dv \\
&\text{letting } n \rightarrow \infty, \\
&\approx \int_0^\infty \frac{(-1)^\ell}{\ell!} f^{(\ell)}(1^-) v^\ell \exp \left[ -\frac{(-1)^\ell \beta_\ell}{(\ell+1)!} v^{\ell+1} \right] dv = \frac{f^{(\ell)}(1^-)}{\beta_\ell} = \frac{f^{(\ell)}(1^-)}{f^{(\ell)}(1^-) + r^{-(\ell+1)} f^{(\ell)}\left(\left(\frac{1}{r}\right)^-\right)}, \quad (40)
\end{aligned}$$

as  $n \rightarrow \infty$  at rate  $O(\kappa_2(f) \cdot n^{-(\ell+2)/(\ell+1)})$ .

For the general case of  $\mathcal{Y} = \{y_1, y_2\}$ , as in the proof of Theorem 4.4, using the transformation  $\phi(x) = (x - y_1)/(y_2 - y_1)$  and replacing  $f(x)$  by  $g(x)$  in Equation (40), the desired result follows. ■